THE UNIVERSITY OF ALBERTA

THE NUMERICAL SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS BY CLOSED DIFFERENCE METHODS

by

WILLIAM ALLAN MURRAY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

EDMONTON, ALBERTA

OCTOBER, 1962

ABSTRACT

The availability of powerful digital computers has stimulated new interest in solving partial differential equations numerically. Closed difference methods are discussed in this thesis; by closed we mean that the solution is completed in a predetermined number of steps.

For completeness, the difference operators needed are defined and difference series for derivatives are derived in this thesis. Before employing them in the numerical solution of partial differential equations, we illustrate their use in the simpler context of ordinary differential equations.

The analytic properties of partial differential equations are discussed briefly since these properties help to determine what numerical technique should be used in a given problem. The difference analogue of a partial differential equation has, however, its own properties and difficulties - properties and difficulties not encountered in the analytic solution. An eminent example of this type of difficulty is the phenomenon of instability: it is possible to translate a well-posed partial differential equation problem into a badly-posed difference problem. This phenomenon is discussed briefly in general terms and more fully in the case of the heat equation and the equation $\frac{\partial^2 U^5}{\partial x^2} = \frac{\partial U}{\partial t}$. The latter is investigated in detail; the analytic solution is obtained and the numerical solution undertaken.

Matrix methods for solving linear elliptic systems are discussed in detail. These methods are only valid if the boundary is a rectangle with sides parallel to the axes, but a procedure for handling non-rectangular boundaries which retains the best qualities of closed methods is suggested. Means by which singularities in boundary data are handled are also considered.

ACKNOWLEDGEMENTS

I wish to thank Professor J. McNamee for his help and guidance. I would also like to acknowledge the consideration shown me by Professor R. V. Andree of the University of Oklahoma and by the Computing Centre of the University of Alberta.

TABLE OF CONTENTS

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

In the past twenty years the fields of physics, chemistry and engineering have called more and more on mathematics for solutions of differential equations and, particularly in recent years, for the solution of partial differential equations. The call has, to some extent, been answered by the development of numerical techniques. Since about 1950, the advent of electronic digital computers has vastly increased our ability to handle many such problems in a reasonable period of time [1].

Accompanying the technological advances in machine design there has been a flood of new methods and techniques. Of these latter, the closed, as opposed to the multiple approximation or relaxation, methods seem particularly interesting since they require little input/output and simpler programs and can be stated in a more concise form. The last point is demonstrated in Chapter V where the boundary value problem is discussed in terms of the solution of algebraic equations.

In essence, the numerical solution of differential equations is simple to describe: the differential equation and its accompanying differential conditions are converted to a difference equation and difference conditions. The difference problem is then susceptible of a numerical solution at discrete points in the region of interest.

This rather general technique will be fully illustrated in later chapters in reference to a particular class of equations, namely partial differential equations of second order in one dependent and two independent variables having real coefficients. Such equations can be classified in

one of three types (Chapter III) and each has a form of numerical solution suited to it. The details of the algorithm for finding the solution, of course, depends on the particular equation and its conditions. For example, the algorithm for solving a boundary value problem is quite different if the derivative rather than the function is specified on the boundary.

The analysis of the errors involved in the approximation by a difference system* of the differential system and of the errors accrued while solving the difference system is extremely difficult and far from complete. This problem has been complicated by the appearance of disturbances of a nature difficult to detect and connect, but very destructive in effect (Chapters II and IV).

The validity of the numerical procedures discussed in this thesis and in the literature rests on the existence and uniqueness theorems of pure mathematics. The existence problem has received much study, with great success in the case of ordinary differential equations and limited success in partials. In many cases we must assume that the problem is well posed (Chapter III) and in fact has a solution. In differential systems describing the physical world this is generally true. That the numerical solution obtained at discrete points in space can be made to conform more and more closely to the actual solution, provided the points are taken closer together, is assumed. This matter is discussed further in Chapter II. Where comparison of the analytic solution and the numerical solution has been possible the results have been encouraging, but as yet little study of other than an empirical nature has been carried out.

---

* By system is meant equation with conditions (boundary or initial).

In what follows, some of the closed difference methods of solving partial difference systems in one dependent and two independent variables will be considered. In Chapter III the common difference operators and their relationships to derivatives will be defined. Since, in many respects, the difference methods for solving ordinary and partial differential systems are quite similar, some consideration will be given to the former in Chapter II. Chapter III is devoted to a brief study of the analytic nature of partial differential equations in order to establish terminology. In Chapter II there is a very brief examination of parabolic systems. Elliptic equations are treated in some detail in Chapter V; closed, matrix methods of solution of elliptic systems are described and a method for handling irregular boundaries is discussed. The analytic solution to the non-linear parabolic equation of Chapter IV is given in Appendix I. It was hoped that a comparison of the analytic and numerical solutions of this equation could have been made, but due to lack of time this was not possible.

Most chapters are preceded by a short introduction.

CHAPTER II

DIFFERENCE OPERATORS AND ORDINARY DIFFERENTIAL EQUATIONS

§2.1. Difference Operators:

The definition of the derivative of a function, $f(x)$, namely

$$\frac{df}{dx} = \lim_{h \to 0} \left(\frac{f(x+h) - f(x)}{h} \right) ,$$

suggests the use of expansions in differences of a function to represent derivatives.

Suppose we are given the sequence of $n+1$ ordinates,

$$f(x_0),\ f(x_1),\ \ldots,\ f(x_n) \ ;$$

corresponding to the sequence of $n+1$ abscissae,

$$x_0,\ x_1,\ \ldots,\ x_n \ ;$$

where, for h a fixed interval,

$$x_i = x_o + ih, \quad i = 0, 1, 2, \ldots, n \ .$$

For convenience, write

$$f(x_i) = f_i \ .$$

The operators, $E, \Delta, \nabla, \delta, \mu, D$; following the notation of Sheppard, are defined as follows:

(2.1)	forward shift,	$Ef_i = f_{i+1}$,
(2.2)	forward difference,	$\Delta f_i = f_{i+1} - f_i$,
(2.3)	backward difference,	$\nabla f_i = f_i - f_{i-1}$,
(2.4)	central difference,	$\delta f_i = f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}$,

(2.5) averaging operation, $\mu f_i = \frac{1}{2}(f_{i+\frac{1}{2}} + f_{i-\frac{1}{2}})$,

(2.6) differentiation operation, $Df_i = (\frac{df}{dx})_{x=x_i}$.

Employing these symbols as operators, most of the common formulae for interpolation, extrapolation, numerical integration and differentiation can be derived [2]. In this thesis we are to be concerned with differential systems and the derivation of appropriate differential difference formulae will be treated in the next section. First, some preliminary results will be obtained. In what follows and in the next section the quantities $E, \Delta, \nabla, \delta, \mu, D,$ will be treated formally. Powers indicate repeated application of the operator, as for example

$$\Delta^2 f_i = \Delta(\Delta f_i) = \Delta(f_{i+1} - f_i) = f_{i+2} - 2f_{i+1} + f_i \quad .$$

From Taylor's theorem

$$\begin{aligned} f_{i+1} = Ef_i &= f_i + hD\, f_i + \frac{h^2}{2!} D^2 f_i + \ldots \\ &= (1 + h\,D + \frac{(hD)^2}{2!} + \ldots)\, f_i \\ &= (e^{hD})\, f_i \quad . \end{aligned}$$

Thus we can write

(2.7) $E = e^{hD}$.

From (2.1), (2.2) and (2.7)

(2.8) $$\begin{aligned} \Delta &= E - 1 \\ &= e^{hD} - 1 \quad . \end{aligned}$$

Using (2.1), (2.3) and (2.7) we obtain

$$(2.9) \qquad \nabla = 1 - E^{-1}$$

$$= 1 - e^{-hD} \; .$$

Combining (2.1) and (2.4) we have

$$\delta = E^{\frac{1}{2}} - E^{-\frac{1}{2}} \; ,$$

and with (2.7) is derived

$$(2.10) \qquad \delta = e^{\frac{1}{2}hD} - e^{-\frac{1}{2}hD}$$

$$= 2 \sinh \tfrac{1}{2}hD \; .$$

From (2.1), (2.5) and (2.7)

$$(2.11) \qquad \mu = \tfrac{1}{2}(E^{\frac{1}{2}} + E^{-\frac{1}{2}})$$

$$= \tfrac{1}{2}(e^{\frac{1}{2}hD} + e^{-\frac{1}{2}hD})$$

$$= \cosh \tfrac{1}{2}hD \; ,$$

Since

$$\cosh^2 \tfrac{1}{2}hD - \sinh^2 \tfrac{1}{2}hD = 1 \; ,$$

we have, from (2.10) and (2.11),

$$(2.12) \qquad \mu = (1 + \tfrac{1}{4}\delta^2)^{\frac{1}{2}} \; .$$

Many other relations of this kind can be derived [2], but these will be sufficient for our purposes.

§2.2. Derivative-difference Relations:

§2.2.1. Forward difference formula:

If we write (2.8) as

$$e^{hD} = 1 + \Delta$$

and take the natural logarithm of both sides we have

$$h\,D = \ln\,(1 + \Delta)\ ,$$

or, for integer power, m,

$$h^m D^m = [\ln\,(1 + \Delta)]^m \tag{2.13}$$

$$= (\Delta - \tfrac{1}{2}\Delta^2 + \tfrac{1}{3}\Delta^3 - \dots)^m \quad .$$

§2.2.2. <u>Backward difference formula</u>:

(2.9) can be written as

$$e^{hD} = (1 - \nabla)^{-1} \quad .$$

Taking the log of both sides and expanding we obtain

$$h^m D^m = (\nabla + \tfrac{1}{2}\nabla^2 + \tfrac{1}{3}\nabla^3 + \dots)^m \quad . \tag{2.14}$$

§2.2.3. <u>Central difference formulae</u>:

Rewrite (2.10) in the form

$$hD = 2\sinh^{-1}\delta/2 \quad .$$

Divide by δ and raise to power m to obtain

$$\left(\frac{hD}{\delta}\right)^m = \left(\frac{\sinh^{-1}\delta/2}{\delta/2}\right)^m \quad .$$

Now

$$\sinh^{-1} x = \int_0^x (1 + t^2)^{-\frac{1}{2}}\,dt$$

$$= \int_0^x (1 - \tfrac{1}{2}t^2 + \tfrac{3}{8}t^4 - \dots)\,dt$$

after formal application of the binomial theorem. Hence,

$$\frac{\sinh^{-1}x}{x} = 1 - \frac{1}{6}x^2 + \frac{3}{40}x^4 - \dots ,$$

and so

$$(2.15) \qquad h^m D^m = \delta^m \left(1 - \frac{\delta^2}{24} + \frac{3}{640}\delta^4 - \dots\right)^m .$$

In particular, for $m = 2$

$$(2.16) \qquad h^2 D^2 = \delta^2 \left(1 - \frac{\delta^2}{12} + \frac{\delta^4}{90} - \dots\right) .$$

It should be noted that for m odd the right hand side of (2.15) contains only odd powers of δ and for m even only even powers. From the definition, (2.4), of δ it can be seen that odd powers of δ involve mid-tabular points, $i \pm 1/2$, $i \pm 3/2$, ..., which, in general, are not known though they can be obtained, with difficulty, by interpolation. A more useful odd derivative central difference formula is as follows:

Multiply the right side of (2.15) above and below by μ^m to get

$$(h\ D)^m = (\mu\delta)^m\ [\mu^{-1}(1 - \frac{\delta^2}{24} + \frac{3}{640}\delta^4 - \dots]^m ,$$

but from (2.12)

$$\mu^{-1} = \left(1 + \frac{\delta^2}{4}\right)^{-\frac{1}{2}}$$
$$= 1 - \frac{\delta^2}{8} + \frac{3}{120}\delta^4 - \dots .$$

Hence, after multiplying the two series together we have

$$(2.17) \qquad h^m D^m = (\mu\delta)^m\ (1 - \frac{1}{6}\delta^2 + \frac{1}{30}\delta^4 - \dots)^m .$$

This form involves ordinates at tabular points only.

For a well behaved function it is preferable to use the central difference formulae where possible since tabular values on both sides of the point of application of the operator are employed.

§2.3. Discussion of Difference Formulae:

To provide justification for the above formulae has proven extremely difficult. Most arguments in support of difference approximations are of an empirical nature and a great number of problems of varying difficulty have been solved analytically and numerically to allow comparison of solutions. The results of such comparisons indicate that we may use difference methods with some confidence.

Two schools of thought, in the matter of difference approximations, stand out. Southwell [3] advocates a crude difference approximation and the use of a very small interval, making the interval smaller when there is any doubt as to the validity of the results, arguing that the discrepancy between

$$\frac{f(x+h) - f(x)}{h} \quad \text{and} \quad \frac{df}{dx}$$

may be made as small as we please provided h is taken small enough. Though this argument seems reasonable it would be difficult to sustain and in addition creates many practical difficulties as we will see in Chapter III. Later work by Fox [4], [5] has shown that very high accuracy can be obtained with a coarse interval if a higher-difference correction to the initially crude difference approximation is introduced. An example of the latter will be given later in this chapter. From a practical point of view the value of the difference correction is undoubted, but again no rigorous justification has been given.

It may be useful at this point to mention briefly some of the theorems of the theory of approximations which are relevant here and to explain their inadequacy in the present context.

The approximation theorems of Weierstrass are of fundamental importance in numerical analysis in as much as they prove that we can approximate a continuous function to any desired degree of accuracy by polynomials or finite trigonometric sums. The theorems, however, are of little practical value since they do not provide an algorithm for actually obtaining the polynomial or trigonometric sum. By accepting a loss of generality results of a more useful nature can be obtained. A proof due to Bernstein is constructive, but the Bernstein polynomials converge so slowly to the function being approximated that little practical use can be made of them.

The basis of the formulae of §2.2 and §2.3 was (2.7), namely

$$f(x+h) = e^{hD} \; f(x)$$

without a remainder term, exact for a limited class of functions. In using these formulae we, in most cases, retain only the first few terms thereby incurring truncation errors. Most analyses of truncation error involve some reference to the remainder term of the Taylor's series (see any standard text where interpolation formulae are derived, for example ref. 6). Hartree [6] shows that (2.7) can be applied exactly to polynomials, exponentials (with linear exponents) products and sums of products of exponentials and polynomials.

§2.4. Difference Equations as Approximations to Ordinary Differential Equations:

§2.4.1. Theoretical Considerations:

Although this thesis is to be mainly concerned with the solving of partial differential systems by means of difference approximations, a brief inspection of methods applicable to ordinary differential systems is helpful since the generalization of methods for O.D.E's* to two or more independent variables is not difficult - formally at least.

Suppose we are given the ordinary differential equation in the dependent variable y and independent variable x

$$f(x,y,y',y'', \ldots ,y^n) = 0 \tag{2.18}$$

with appropriate conditions. Primes denote differentiation with respect to x. We require that (2.18) be solvable for the highest derivative. It is required to find $y = y(x)$.

The theory of O.D.E's* yields an extremely valuable theorem establishing the existence and uniqueness of the solution to the set of first order equations

$$\begin{aligned} y_1' &= f_1(x,y_1,y_2,\ldots,y_n) \,, \\ y_2' &= f_2(x,y_1,y_2,\ldots,y_n) \,, \\ &\ldots\ldots\ldots\ldots\ldots, \\ y_n' &= f_n(x,y_1,y_2,\ldots,y_n) \end{aligned} \tag{2.19}$$

with initial conditions

---

* O.D.E. - ordinary differential equation.

$$(2.20) \qquad y_1(x_o) = y_1^o\ ,\quad y_2(x_o) = y_2^o,\ \ldots,\quad y_n(x_o) = y_n^o\ .$$

That (2.18) is equivalent to the set (2.19) is seen as follows: solve (2.18) for the derivative of highest order as

$$y^n = g(x, y, y', \ldots, y^{n-1})\ .$$

Let

$$y_1 = y,\ \ y_2 = y',\ \ldots,\ \ y_n = y^{n-1}\ .$$

Hence we have the set of equations

$$y_1' = y_2\ ,$$

$$y_2' = y_3\ ,$$

$$\ldots\ ,$$

$$y_{n-1}' = y_n\ ,$$

$$y_n' = g(x, y_1, y_2, \ldots, y_n)\ ;$$

which is a special case of (2.18).

We may state the theorem as follows: "Given a differential system (2.19), (2.20); if two positive numbers a and b can be found such that in the domain D defined by the inequalities

$$x_o \leq x \leq x_o + a\ ;$$

$$y_1^o - b \leq y_1 \leq y_1^o + b,\quad y_2^o - b \leq y_2 \leq y_2^o + b\ \ldots,$$

$$y_n^o - b \leq y_n \leq y_n^o + b$$

the functions $f_1, f_2, \ldots, f_n$ are continuous and satisfy Lipschitz conditions in the variables $y_1, y_2, \ldots, y_n$ within the ranges, then a

positive number $\delta(\leq a)$ can be determined such that in the interval $(x_o, x_o+\delta)$ the system possesses one and only one solution".

The proof of this theorem will not be given here. Similar theorems for boundary conditions can also be found [7].

Although the existence and uniqueness of the solution to a given O.D.E. and P.D.E.* system may have been established by the general theorems, obtaining the solution in analytic form is another matter. In addition, even if such a solution can be found it is often extremely difficult to obtain numerical values from it (see for example Chapter IV).

§2.4.2. Numerical Considerations:

In general terms, the numerical solution of an O.D.E. system consists of replacing derivatives, where they appear in the equation and conditions, by a few terms of an expression such as are given in §2.2 and solving, if possible, the difference system so obtained for the numerical values of the function at tabular intervals. In non-linear systems approximations to powers of derivatives must be used. Later in this chapter we will see that the difference system should not, in general, be of higher order than the differential system. Additional terms of the difference formula may be included as a correction after a first approximation to the solution has been obtained.

In exemplification of this statement let us consider

$$\frac{dy}{dx} = g(x,y) ,$$

$$y(x_o) = a \quad .$$

---

* P.D.E. - partial differential equation.

Writing

$$y(x_o + ih) = y_i ,$$

$$x_o + ih = x_i ,$$

replace dy/dx by the first term of (2.13). The difference system to be solved is

$$y_{i+1} = y_i + g(x_i,y_i) ,$$

$$y_o = a , \quad i = 0, 1, 2, \ldots .$$

After obtaining the first approximation, ${}^{(1)}y_i$, we may include further terms of (2.13) and solve

$${}^{(2)}y_{i+1} = {}^{(2)}y_i + g(x_i, {}^{(2)}y_i) + \frac{1}{2}\Delta^2 \, {}^{(1)}y_i - \frac{1}{3}\Delta^3 \, {}^{(1)}y_i + \ldots ,$$

$${}^{(2)}y_o = a , \quad i = 0,1,2,\ldots .$$

Generally only one or two correction terms are used.

§2.4.3. <u>The Problem of Stability</u>:

In §2.3 we mentioned one of the sources of error incurred in the numerical solution of O.D.E. and P.D.E. systems, that is, the truncation of the difference series. Despite the difficulty of analysing such error we can place confidence in difference corrections and the use of small intervals. Inherent in all numerical work there are errors due to rounding of figures to a limited number of places and due to loss of significant figures resulting from arithmetic operations. Detailed analyses of the latter has been carried out for some of the well known procedures in numerical analysis [8], [9]. Except where such procedures are used, the error analysis of each problem depends on the characteristics of the problem and more basic techniques must be used [6].

In addition to the above mentioned, a deviation between the actual solution to an O.D.E. or P.D.E. system and the associated numerical solution often occurs which defies correction and which exceeds the most generous round-off-error bounds. This is the problem of stability. No detailed account of instability in the solution of O.D.E. systems will be given here. Stability in P.D.E. methods will be mentioned later (Chapter III). A simple example will serve to indicate the nature of instability in the solution of O.D.E. systems.

Consider the second order system

$$y'' - y = 0 ,$$
$$(2.21) \qquad y(o) = 1, \quad y(h) = e^{-h}; \quad h > 0 .$$

The analytic solution is

$$y(x) = e^{-x} .$$

Using the first term of equation (2.16), we obtain the difference system

$$y_{n+1} = (2+h^2)\, y_n - y_{n-1} ,$$
$$(2.22) \qquad y_o = 1, \quad y_1 = e^{-h}, \quad n = 1, 2, 3, \dots ;$$

where $x_n = nh$, $y_n = y(x_n)$.

Table I shows the solution of (2.22) for $h = 0.5, 0.25, 0.1$. The figures in brackets are the differences between e^{-nh} and the numerical solution in units of the last figure retained.

Note that the error increases alarmingly and rapidly and that the solution eventually becomes negative, even when we decrease h to 0.1. To discover the cause of this behavior consider the difference equation.

TABLE I

SOLUTION OF $y(x+h) = (2+h^2)\, y(x) - y(x-h)$

h = 0.5		h = 0.25		h = 0.1	
x	y(x)	x	y(x)	x	y(x)
0	1.00000	0	1.00000	0	1.00000
0.5	.60653	0.25	.77880	0.1	.90484
1.0	.36469 (319)	0.50	.60627 (26)	0.2	.81873 (0)
1.5	.21402 (915)	0.75	.47163 (74)	0.3	.74081 (1)
2.0	.11685 (1848)	1.00	.36646 (142)	0.4	.67030 (2)
2.5	.04889 (3320)	1.25	.28419 (231)	0.5	.60649 (4)
3.0	-.00685	1.50	.21968 (345)	. .	. . . .
		. . .	. . . .	1.0	.36762 (26)
		3.25	.01293 (2585)	. .	. . . .
		3.50	-.00314	3.0	.04638 (361)
				. .	. . . .
				4.4	-.00175

Try a solution of the form

$$y_n = c^n \ , \quad c \ \text{ a constant} \ .$$

If we substitute this in the difference equation of (2.22) and divide through by c^{n-1} we obtain

$$c^2 - (2+h^2)c + 1 = 0$$

which has the solutions

$$c_1 = 1 + \tfrac{1}{2}h^2 + h\sqrt{1+\tfrac{1}{4}h^2} \quad ,$$

$$c_2 = 1 + \tfrac{1}{2}h^2 - h\sqrt{1+\tfrac{1}{4}h^2} \quad .$$

Since

$$c_1 c_2 = 1 \ ,$$

then

$$c_2 = c_1^{-1} \quad .$$

Therefore the general solution is

$$y_n = A\, c^n + Bc^{-n} \quad ,$$

where

$$c = c_1 > 1 \quad .$$

Applying the data of (2.22) we have

$$A + B = 1$$

and $$Ac + Bc^{-1} = e^{-h} \quad .$$

Hence $$B = 1 - A$$

and

$$A = \frac{e^{-h} - c^{-1}}{c - c^{-1}} \quad .$$

The source of our trouble is now apparent. The coefficient, A, of the increasing solution to the difference equation is not zero and despite the smallness of h this solution will disturb and finally dominate the solution we are seeking.

In essence, therefore, instability consists of the introduction of large unwanted solutions to the difference system. An instable system may be created in many ways: by converting a stable differential system

to the difference system, by using a difference system of higher order than the differential system, and by trying to solve an inherently unstable differential equation.

It should be noted that "marching" techniques, which must be used for initial value problems, (a crude version of marching was used in the above example) can be expected to exhibit instability since they "precede" the data and "predict" the value of the solution at the next tabular point. Generally the numerical methods applicable to boundary value problems are stable provided the O.D.E. system is stable and the order of the difference equation does not exceed that of the differential equation.

CHAPTER III

SOME GENERAL ASPECTS OF PARTIAL DIFFERENTIAL EQUATIONS

§3.1. Introduction:

Before discussing in detail some numerical procedures for solving P.D.E. systems, let us consider these equations briefly in a more general way in order to establish terminology and concepts which will be of use in later chapters. The following is not meant to be complete or rigorous. Some material which is found in the literature [10,11] is included here for completeness of exposition.

§3.2. Definition of a Partial Differential Equation, Existence and Uniqueness:

A P.D.E. of order k in the n independent variables

$$x_1, x_2, \ldots, x_n$$

and one dependent variable

$$U = U(x_1, x_2, \ldots, x_n)$$

is a relation of the form

$$F\left(x_i, \frac{\partial^\ell U}{\partial x_r \, \partial x_s \ldots \partial x_t} \right) = 0, \qquad \begin{array}{l} i,r,s,t = 1,2,\ldots,n, \\ \ell = 0,1,2,\ldots,k; \end{array} \tag{3.1}$$

where F is a function of the independent variables, the dependent variable and the partial derivatives (of all orders up to and including k) of the dependent variable.

The <u>order</u> of the P.D.E. is the order of the highest derivatives (k).

The equation is linear if F is linear in U and the partial derivatives of U with coefficients depending only on the independent variables.

If F is linear in the kth derivatives with coefficients depending upon $x_1, x_2, \ldots, x_n$, U and perhaps the partial derivatives of U up to order $k - 1$, then the equation is quasi-linear, otherwise it is non-linear.

The theory of P.D.E 's is not as complete as that of O.D.E 's. There is no general theorem of uniqueness and existence comparable to that of §2.4.1 although some success has been achieved in special classes of equations [10]. The theorem of Cauchy and Kowalewsky is one of the most general though it imposes severe restrictions on the P.D.E. Like the theorem of §2.4.1 the Cauchy-Kowalewsky proof [12] applies to a set of P.D.E 's in the form

$$\frac{\partial^k U^i}{\partial x_1^{\,k}} = f_i\left(x_1, x_2, \ldots, x_n, \frac{\partial U^1}{\partial x_1}, \ldots, \frac{\partial^k U^m}{\partial x_n^{\,k}} \right),$$

$$i = 1, 2, \ldots, n \; .$$

This is the normal form. The f_i must be analytic functions of the n independent variables and the partial derivatives of the m dependent variables. Initial data is given on the initial plane $x_1 = 0$ and is the form

$$U^i\,(0, x_2, \ldots, x_n) = \varphi_{i,o}\,(x_2, x_3, \ldots, x_n)\,,$$

$$\frac{\partial^t U^i}{\partial x^t_{\;1}} = \varphi_{i,t}\,(x_2, x_3, \ldots, x_n)\,,$$

where

$$i = 1, 2, \ldots, m ,$$

$$t = 1, 2, \ldots, k-1$$

and the functions, φ, must be analytic. Under these conditions a unique solution,

$$U^1, U^2, \ldots, U^m ,$$

exists.

For the purposes of this discussion we will consider only linear and quasi-linear second order equations in two independent and one dependent variable, that is, equations of the form

$$a\, U_{xx} + 2b\, U_{xy} + c\, U_{yy} + d = 0 \quad , \tag{3.2}$$

where a, b, c, d are real functions of x, y, U, U_x, U_y (subscripts denote partial differentiation). The quantities x, y do not necessarily denote Cartesian co-ordinates.

§3.3. Solution Using Initial Data on a Given Curve:

If O.D.E. theory, when initial values are given, we are often able to construct a solution in the form of a Taylor's series about the initial point. For example, given the system

$$y'' = f(x,y,y'),$$
$$y(o) = y_o,\ y'(o) = p ;$$

we can obtain $y''(o)$ from the equation. For $y'''(o)$ we have

$$y''' = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\, y' + \frac{\partial f}{\partial y'}\, y'' \quad ,$$

and so on. A similar technique can be applied to the partial differential

initial value problem, where now the function and its partial derivatives of first order are given along some initial curve.

Let

$$U_x = p,\ U_y = q,\ U_{xx} = r,\ U_{xy} = s,\ U_{yy} = t\ .$$

Then (3.2) can be written as

$$ar + 2bs + ct + d = 0\ .$$

Let ℓ be some parameter on the curve L in x, y, U - space on which we are given the data

$$(3.3) \qquad x = x(\ell),\quad y = y(\ell),\quad U = U(\ell),\quad p = p(\ell),\quad q = q(\ell)\ .$$

Let us attempt to construct a solution to the P.D.E. system (3.2), (3.3) by a Taylor's series expansion about some point (x_o, y_o, U_o) on L:

$$U(x,y) = U_o + (x-x_o)p_o + (y-y_o)q_o + (x-x_o)^2 r_o$$

$$+ (x-x_o)(y-y_o)s_o + (y-y_o)^2 t_o + \dots\ .$$

For this we must find the higher order derivatives of U at (x_o, y_o) (we assume that derivatives of all orders of U do, in fact, exist).

Now

$$dp = \frac{\partial p}{\partial x}dx + \frac{\partial p}{\partial y}dy$$

$$= r\,dx + s\,dy\ .$$

Thus

$$\frac{dp}{d\ell} = r\frac{dx}{d\ell} + s\frac{dy}{d\ell}\ .$$

Similarly

$$\frac{dq}{d\ell} = s\,\frac{dx}{d\ell} + t\,\frac{dy}{d\ell} \quad .$$

These equations, together with (3.2), give us three equations in the three unknowns r, s, t; namely:

$$(3.4) \qquad Az = f \quad ,$$

where

$$A = \begin{bmatrix} \frac{dx}{d\ell} & \frac{dy}{d\ell} & 0 \\ 0 & \frac{dx}{d\ell} & \frac{dy}{d\ell} \\ a & 2b & c \end{bmatrix} \quad ,$$

$$z = \begin{bmatrix} r \\ s \\ t \end{bmatrix} \quad \text{and} \quad f = \begin{bmatrix} \frac{dp}{d\ell} \\ \frac{dq}{d\ell} \\ -d \end{bmatrix} \quad .$$

We can determine further derivatives of U by differentiating successively (3.4). For example

$$A\,\frac{\partial z}{\partial x} = \frac{\partial f}{\partial x} - \frac{\partial A}{\partial x}\,z \quad , \quad A\,\frac{\partial z}{\partial y} = \frac{\partial f}{\partial y} - \frac{\partial A}{\partial y}\,z$$

from which we can determine

$$\frac{\partial z}{\partial x} = \begin{bmatrix} U_{xxx} \\ U_{xxy} \\ U_{xyy} \end{bmatrix} \quad , \quad \frac{\partial z}{\partial y} = \begin{bmatrix} U_{xxy} \\ U_{xyy} \\ U_{yyy} \end{bmatrix} \quad .$$

Obviously, we can continue this procedure indefinitely unless we are

stopped at the beginning. Suppose that the determinant of A is zero at the point on L that we are considering, that is, suppose that

$$|A| = a\left(\frac{dy}{d\ell}\right)^2 - 2b\,\frac{dx}{d\ell}\,\frac{dy}{d\ell} + c\left(\frac{dx}{d\ell}\right)^2 = 0 .$$

This condition can be written as

$$(3.5) \qquad a(y')^2 - 2by' + c = 0 , \quad y' = \frac{dy}{dx} ,$$

which yields two values of the slope, $y'=\frac{dy}{dx}$, namely

$$y' = \frac{b \pm (\Delta)^{\frac{1}{2}}}{a} , \quad \Delta = b^2 - ac .$$

Thus, depending on whether

$$\Delta > 0, = 0, < 0 ,$$

equation (3.5) defines two families of curves in the x-y plane, one family or no real curves respectively. Notice that these curves depend only on the coefficients, a, b, c, of equation (3.2). These are called the <u>characteristic curves</u> of the equation.

Two points should be made clear: our Taylor's series method fails if we proceed in either of the directions defined by (3.5), and we cannot obtain a solution by this method no matter what direction we go if the slope of L coincides with either of these directions. The characteristic curves (if real) seem to form a natural boundary across which the solution <u>may</u> be discontinuous.

It must not be supposed that data at one point serves to determine the solution throughout the x,y-plane. Since the characteristics bound the region of validity of the Taylor's series solution about a point on the initial curve, all points on this curve must be

used to obtain the entire solution.

§3.4. Classification of Partial Differential Equations:

A further property of the discriminant,

$$\Delta = b^2 - ac \quad ,$$

is that its sign determines to which of the three standard forms (hyperbolic, parabolic or elliptic) equations of type (3.2) can be reduced.

Let us transform (3.2), by a change of independent variables, into a form in which the second order, cross derivative does not appear, that is,

$$\gamma\, U_{\zeta\zeta} + \delta\, U_{\eta\eta} + d' = 0 \quad ,$$

where γ, δ, d' are functions of $\zeta, \eta, U, U_\zeta, U_\eta$.

Let

$$\zeta_x = \varphi b \ , \quad \eta_x = \psi b, \tag{3.6}$$

$$\zeta_y = \varphi(\lambda_1 - a) \ , \quad \eta_y = \psi\,(\lambda_2 - a) \ ,$$

where λ_1 and λ_2 are the roots of

$$\lambda^2 - (a+c)\,\lambda - \Delta = 0 \quad , \tag{3.7}$$

and φ, ψ are chosen so that

$$\zeta_{xy} = \zeta_{yx} \ , \quad \eta_{xy} = \eta_{yx} \quad .$$

Now

$$\frac{\partial U}{\partial x} = \zeta_x \frac{\partial U}{\partial \zeta} + \eta_x \frac{\partial U}{\partial \eta} \quad .$$

Hence,

$$\frac{\partial^2 U}{\partial x^2} = (\zeta_x)^2 \; U_{\zeta\zeta} + 2\zeta_x\eta_x \, U_{\zeta\eta} + (\eta_x)^2 \, U_{\eta\eta}$$

$$= (\varphi b)^2 \, U_{\zeta\zeta} + 2\varphi\psi b^2 \, U_{\zeta\eta} + (\psi b)^2 U_{\eta\eta}$$

and

$$\frac{\partial^2 U}{\partial x \partial y} = \zeta_x \; \zeta_y \; U_{\zeta\zeta} + (\zeta_y\eta_x + \zeta_x\eta_y) \; U_{\zeta\eta} + \eta_x \; \eta_y \; U_{\eta\eta}$$

$$= \varphi^2 b(\lambda_1 - a) + \varphi\psi b(\lambda_1 - 2a + \lambda_2) \; U_{\zeta\eta} + \psi^2 b(\lambda_2 - a) \; U_{\eta\eta} \quad .$$

Similarly

$$\frac{\partial^2 U}{\partial y^2} = \varphi^2(\lambda_1 - a)^2 \; U_{\zeta\zeta} + 2\varphi\psi \; (\lambda_1 - a)(\lambda_2 - a) \; U_{\zeta\eta} + \psi^2(\lambda_2 - a)^2 \; U_{\eta\eta} \; .$$

Substitute these last relations into (3.2).

The coefficient of $U_{\zeta\eta}$ is

$$\mu = 2\varphi\psi \; [ab^2 + b^2(\lambda_1 - 2a + \lambda_2) + c(\lambda_1 - a)(\lambda_2 - a)]$$

$$= 2\varphi\psi \; [c\lambda_1\lambda_2 + (\lambda_1 + \lambda_2)(b^2 - ac) + a(ac - b^2) \;] \quad ,$$

but from (3.6)

$$\lambda_1 \; \lambda_2 = -\Delta \, , \quad \lambda_1 + \lambda_2 = a + c \quad .$$

Hence,

$$\mu = 2\varphi\psi\Delta \; [-c + a + c - a]$$

$$= 0 \quad .$$

The coefficient of $U_{\zeta\zeta}$ is

$$\gamma = \varphi^2 \; [ab^2 + 2b^2 \; (\lambda_1 - a) + c(\lambda_1 - a)^2] \quad .$$

The coefficient of $U_{\eta\eta}$ is

$$\delta = \psi^2 \, [ab^2 + 2b^2(\lambda_2 - a) + c(\lambda_2 - a)^2] \quad .$$

Ignoring for the moment the distinction between λ_1 and λ_2, we can write γ / φ^2 and δ / ψ^2 as

$$\epsilon = ab^2 + 2b^2\,(\lambda - a) + c(\lambda - a)^2 \; .$$

From (3.7) we have

$$b^2 = (\lambda - a)(\lambda - c) \quad .$$

Replace b^2 in the first term of the expression for ϵ by the above expression and we obtain

$$\begin{aligned}
\epsilon &= a(\lambda - a)(\lambda - c) + 2b^2(\lambda - a) + c(\lambda - a)^2 \\
&= (\lambda - a)\;[\lambda(a+c) + 2(b^2 - ac)] \\
&= (\lambda - a)\;[\lambda(\lambda_1 + \lambda_2) - 2\,\lambda_1\lambda_2] \quad .
\end{aligned}$$

Thus

$$\gamma / \varphi^2 = (\lambda_1 - a)\;\lambda_1\;(\lambda_1 - \lambda_2) \; ,$$

$$\delta / \psi^2 = -(\lambda_2 - a)\;\lambda_2\;(\lambda_1 - \lambda_2)$$

and so

$$\frac{\gamma\delta}{\varphi^2\psi^2} = -\,(\lambda_1 - a)(\lambda_2 - a)\;\lambda_1\;\lambda_2\;(\lambda_1 - \lambda_2)^2 \quad .$$

Now

$$\begin{aligned}
(\lambda_1 - a)(\lambda_2 - a) &= \lambda_1\lambda_2 - a(\lambda_1 + \lambda_2) + a^2 \\
&= -b^2 + ac - a(a+c) + a^2 \\
&= -b^2 \quad .
\end{aligned}$$

Therefore

$$\frac{\gamma\delta}{\varphi^2\psi^2} = -b^2 (\lambda_1 - \lambda_2)^2 \Delta ,$$

where we choose

$$\lambda_1 = \frac{a+c - [(a+c)^2 + 4\Delta]^{\frac{1}{2}}}{2} ,$$

$$\lambda_2 = \frac{a+c + [(a+c)^2 + 4\Delta]^{\frac{1}{2}}}{2} .$$

There are three cases to consider.

Case (i): $\Delta = 0$.

Hence,

$$\lambda_1 = 0, \quad \lambda_2 = a+c$$

and so

$$\gamma = 0 ,$$

$$\delta = \psi^2 c(a+c)^2 .$$

Case (ii): $\Delta > 0$.

Signs of γ and δ are opposite.

Case (iii): $\Delta < 0$.

Signs of γ and δ are the same.

Therefore, if we write

$$\alpha^2 = \varphi^2 \ |(\lambda_1 - a)\ \lambda_1\ (\lambda_1 - \lambda_2)| ,$$

$$\beta^2 = \psi^2 \ |(\lambda_2 - a)\ \lambda_2\ (\lambda_1 - \lambda_2)|$$

equation (3.2) can be reduced to one of the following standard forms:

$$\psi^2 c(a+c)^2 \ \frac{\partial^2 U}{\partial \eta^2} + d' = 0 , \quad \Delta = 0 ;$$

$$\alpha^2 \frac{\partial^2 U}{\partial \zeta^2} - \beta^2 \frac{\partial^2 U}{\partial \eta^2} + d' = 0 , \quad \Delta > 0 ;$$

$$\alpha^2 \frac{\partial^2 U}{\partial \zeta^2} + \beta^2 \frac{\partial^2 U}{\partial \eta^2} + d' = 0 , \quad \Delta < 0 .$$

By comparing these equations with the algebraic equations

$$y^2 = \text{const.}\ x \ ,$$

$$\alpha^2 x - \beta^2 y^2 = \text{const.},$$

$$\alpha^2 x^2 + \beta^2 y^2 = \text{const.},$$

we name the standard forms parabolic, hyperbolic and elliptic respectively. Notice that parabolic P.D.E 's have one real characteristic curve, hyperbolic equations have two and elliptic equations have none.

As an example consider the equation

$$U_{xx} + y\,U_{yy} = 0 \ .$$

Now

$$\Delta = b^2 - ac = -y \ .$$

Thus, the P.D.E. is hyperbolic in the half plane

$$y < 0 \ ,$$

elliptic in the half plane

$$y > 0 \ ,$$

and parabolic along the line

$$y = 0 \ .$$

The characteristic curves are given by

$$\left(\frac{dy}{dx}\right)^2 + y = 0 \ ,$$

which integrates to

$$y = -\frac{1}{4}(x + e)^2 , \quad e \text{ being a constant.}$$

These curves are real only if $y \leq 0$.

It does not follow that transforming a given P.D.E. to standard form provides the best method for practical solution; indeed, such a transformation may unnecessarily complicate the problem since the given data must also be transferred. The practical difficulty of solving equations (3.6) for $\zeta(x,y)$ and $\eta(x,y)$ and equation (3.5) for the characteristics should not be ignored, since if the equation is not linear this involves finding the solution to the equation as well. Several methods for solving quasi-linear systems in terms of characteristics have been developed [10] which involve the simultaneous determination of ζ, η, U, U_ζ and U_η . Corresponding numerical procedures exist [13].

§3.5. Data for Partial Differential Equations:

In the preceding sections of this chapter we found that hyperbolic, parabolic and elliptic P.D.E 's differed from each other in the possession of characteristic curves. It is reasonable to ask if there are other properties which distinguish these equations from each other.

In a practical sense it is quite meaningless to talk of solving a P.D.E. without making reference to certain auxiliary restrictions on the dependent variable and, perhaps, its derivatives. We will define boundary conditions to mean data given on a closed curve in the plane of the dependent variable, and initial conditions to mean data given on an open curve. We might ask if each type of P.D.E. requires boundary or initial data in order that the problem be well-posed in the sense that the solution should exist, be uniquely determined and depend continuously on the data.

Let us consider a few examples from physics, assuming that P.D.E. systems resulting from the correct application of physical laws represent well-posed problems.

§3.5.1. <u>The wave equation in one space dimension</u>:

The equation of motion of a string of finite length, under uniform tension T, with uniform density ϵ, undergoing small amplitude motion is

$$\frac{\partial^2 U}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} , \quad c^2 = T / \epsilon > 0 ,$$

where x is the space co-ordinate and t is the time co-ordinate. Clearly, the equation is hyperbolic.

Auxiliary conditions might be obtained by fixing the ends of the string and displacing it from the point of rest $(x=0)$ into some configuration, say $f(x)$, then releasing it. If the ends of the string are the points

$$x = a, \quad x = b$$

the auxiliary conditions can be stated as follows:

$$U(a,t) = U(b,t) = 0 ,$$
$$U(x,0) = f(x) ,$$
$$U_t(x,0) = 0 ,$$
$$t \geq 0 , \quad a \leq x \leq b .$$

Conditions such as these are initial since they are given on an open curve defined by

$$x = a, \quad x = b, \quad t = 0 .$$

Hence, we might suspect that hyperbolic equations require initial conditions.

§3.5.2. Poisson's equation:

Suppose we are given a long rectangular box of width a, height b, filled with a dielectric medium. The opposing sides of the box are at potential V_1, V_2, V_3, V_4 respectively and the charge density in the medium is $f(x,y)$. If U is the potential in the box, then

$$U = U(x,y)$$

with no dependence on z. The equation for the potential is

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = f(x,y) \quad .$$

If the box is symmetric about the origin, the auxiliary conditions are

$$U(\tfrac{a}{2}, y) = V_1 \;, \quad U(-\tfrac{a}{2}, y) = V_2 \;,$$

$$U(x, \tfrac{b}{2}) = V_3 \;, \quad U(x, -\tfrac{b}{2}) = V_4 \;.$$

The equation is elliptic and the data is of boundary type. We anticipate that well-posed elliptic problems require boundary data.

§3.5.3. The heat equation:

Let $U(x,t)$ be the temperature at time t and at point x of a narrow, uniform rod lying along the x - axis. If there are no heat sources or sinks in the rod, the equation governing the temperature distribution is

$$a^2 \frac{\partial^2 U}{\partial x^2} = \frac{\partial U}{\partial t} \;, \quad t \geq 0 \;, \quad x \geq 0 \;,$$

where a^2 is the constant of heat diffusivity of the rod. This equation is parabolic.

Suppose the rod is infinitely long and initially has the temperature distribution $f(x)$. Then the appropriate condition is

$$U(x,0) = f(x)$$

which is an initial condition. Hence, parabolic systems appear to be initial.

In summary then, we are led to suspect, by considering physical problems, that well-posed hyperbolic, elliptic and parabolic systems are initial, boundary and initial value problems respectively. Examples can be given [12] to show that applying initial data to elliptic equations or boundary data to hyperbolic or parabolic equations will result in systems which are not well-posed.

§3.6. Remarks on the Numerical Solution of Partial Differential Systems:

§3.6.1. Difference approximations to partial differential systems:

Although the numerical solution of P.D. systems has received much attention since early in the 1940's [1,3,14] (see esp. [1] for a comprehensive bibliography) little of a general nature can be said.

In this section a procedure for reducing equations of type (3.2) to difference equations will be discussed. Essentially, this reduction is the same as that of an O.D. system except that the solution is sought at discrete points in a two-dimensional region rather than on a line.

For the moment, assume that the data curve is a rectangle composed of the axes and the lines $x = 1$, $y = 1$ if it is a boundary problem. If the problem is of initial type, assume the initial curve is an open rectangle bounded on three sides by the axes and the line $x = 1$.

By a linear change of variables we can always obtain this configuration provided the data curve is rectangular. Non-rectangular curves present a difficulty to be discussed later. Let us also assume that we are to find solutions to the initial value problems in the range $0 \leq y \leq 1$.

Let us examine in detail the procedure for setting up and solving the difference system corresponding to the P.D. system

$$aU_{xx} + bU_{xy} + cU_{yy} + d = 0 ,$$

$$0 \leq x \leq 1 , \quad 0 \leq y \leq 1 .$$

If the system is initial, let the data curve be

$$x = 0 , \quad x = 1 , \quad y = 0 .$$

If the system is boundary, consider the data curve

$$x = 0 , \quad x = 1 , \quad y = 0 , \quad y = 1 .$$

A mesh or grid is formed in the region of solution composed of vertical and horizontal lines through the points

$$x_i = ih_x ,$$

$$y_j = jh_y ,$$

where

$$h_x = 1/n , \quad h_y = 1/m$$

and

$$i = 0, 1, 2, \ldots, n ,$$

$$j = 0, 1, 2, \ldots, m .$$

Let

$$U(x_i, y_j) = U_{ij}$$

and let corresponding expressions for a, b, c, d be denoted a_{ij}, b_{ij}, c_{ij}, d_{ij}.

Partial derivatives with respect to x and y are replaced by difference formulae drawn from §2.2 with differences taken in the x and y directions respectively. Subscripts placed on difference operators will denote the direction in which differencing is to occur. The difference equation can now be written down at each point of the region of solution, as for example,

$$a_{ij}\, \delta_x^2\, U_{ij} + b_{ij}\, \mu_x\, \delta_x(\mu_y\, \delta_y\, U_{ij}) + c_{ij}\, \delta_y^2\, U_{ij} + d_{ij} = 0\ .$$

If the equation is not linear we can say little more in general since completion of the algorithm depends on the particular equation being considered. It is common practice to linearize the difference equation by the device of replacing non-linear terms by linear approximations. A crude first approximation to the solution can be obtained in this way. After the first solution is obtained we can linearize the equation by using this estimate in much the same manner as difference corrections are applied. For example, a term such as U_{ij}^3 might be replaced by $({}^{(1)}U_{ij}^2)\ {}^{(2)}U_{ij}$, where ${}^{(1)}U_{ij}$ is the known first estimate and ${}^{(2)}U_{ij}$ is to be found. Difference corrections can be applied as well. An example of the above procedure is given in Chapter IV where the equation

$$\frac{\partial^2 U^5}{\partial x^2} = \frac{\partial U}{\partial t}$$

is considered.

Even if the P.D.E. is linear a complete discussion of the algorithms for solving the P.D. system would necessitate considering all possible data conditions, each requiring the difference equation in a different form. A particular example is given below.

§3.6.2. Example:

To illustrate what may be called the general procedure for setting up the closed difference solution to a P.D. system, let us consider the following example:

Given the linear, hyperbolic P.D.E.

$$aU_{xx} + bU_{xy} + cU_{yy} + dU_x + eU_y + fU + g = 0 \tag{3.8}$$

and the data

(3.8a) $U(0,y) = A(y)$,

(3.8b) $U(1,y) = B(y)$,

(3.8c) $U(x,0) = C(x)$,

(3.8d) $\left(\frac{\partial U}{\partial y}\right)_{y=0} = D(x)$,

$$0 \leq x \leq 1 \ , \quad 0 \leq y \leq 1 \ ,$$

where $a, b, c, d, e, f, g, A, B, C, D$ are well-behaved functions. Assume that the data has no discontinuities at the corners.

We require that the equation be hyperbolic throughout the region of solution and that the line $y = 0$ not be a characteristic. It is assumed that the solution exists, is unique and that all derivatives of U with respect to x and y exist.

Using (3.8), (3.8c) and (3.8d) we can obtain the first row

solution, U_{i1}, by a Taylor's series:

$$U_{i1} = U(x,h_y) = U(x,0) + h_y \left(\frac{\partial U}{\partial y}\right)_{y=0} + h_y^2 \left(\frac{\partial^2 U}{\partial y^2}\right)_{y=0} + \cdots \quad .$$

Write (2.16) in the form

$$h^2 D^2 = \delta^2 + c_1$$

and (2.17) as

$$hD = \mu\delta + c_2$$

$$= \frac{1}{2}(E - E^{-1}) + c_2 \ .$$

From the latter we have

$$h_x h_y U_{xy} = (\mu_x \delta_x + c_{x2})(\mu_y \delta_y U + c_{y2} U)$$

$$= \mu_x \delta_x (\mu_y \delta_y U) + c_3 U \ .$$

The expressions represented by c_1, c_2, c_3 are difference-correction terms.

If we multiply (3.8) by $h_x^2 h_y^2$ and replace derivatives by difference formulae, we obtain

$$ah_y^2(\delta_x^2 U + c_{x1} U) + bh_x h_y[\mu_x \delta_x(\mu_y \delta_y U) + c_3 U] + ch_x^2(\delta_y^2 U + c_{y1} U)$$

$$+ dh_x h_y^2(\mu_x \delta_x U + c_{x2} U) + eh_x^2 h_y(\mu_y \delta_y U + c_{y2} U) + fh_x^2 h_y^2 U + gh_x^2 h_y^2 = 0 \ .$$

Dividing through by h_x^2, taking all correction terms to the right and writing

$$U(x,y) = U_{ij}$$

gives us the difference equation

$$a_{ij}k^2\delta_x^2U_{ij} + b_{ij}k\mu_x\delta_x(\mu_y\delta_yU_{ij}) + c_{ij}\delta_y^2U_{ij} + d_{ij}h_yk\mu_x\delta_xU_{ij}$$

$$+ e_{ij}h_y\mu_y\delta_yU_{ij} + f_{ij}h_y^2U_{ij} + g_{ij}h_y^2 = c_5U_{ij} ,$$

where

$$i = 1, 2, \ldots, n - 1 ,$$

$$j = 1, 2, \ldots, m - 1 ,$$

$$k = h_y/h_x .$$

The given data is

$$U_{oj} = A_j , \quad U_{nj} = B_j , \quad U_{io} = C_i$$

and U_{i1} has been calculated.

Write the difference equation explicitly in terms of the unknown row, $U_{\ell,j+1}$ $(\ell=0,1,2,\ldots,n)$, as follows:

$$-\left(\frac{b_{ij}k}{4}\right)U_{i-1,j+1} + \left(c_{ij}+\frac{e_{ij}h_y}{2}\right)U_{i,j+1} + \left(\frac{b_{ij}k}{4}\right)U_{i+1,j+1}$$

$$= c_5U_{ij} - \left[a_{ij}k^2\delta_x^2U_{ij} + \frac{b_{ij}k}{4}\left(U_{i-1,j-1}-U_{i+1,j-1}\right)\right.$$

$$+ c_{ij}\left(U_{i,j-1}-2U_{ij}\right) + \frac{d_{ij}h_yk}{2}\mu_x\delta_xU_{ij} - \frac{e_{ij}h_y}{2}U_{i,j-1}$$

$$\left. + h_y^2\left(f_{ij}U_{ij} + g_{ij}\right)\right] .$$

To simplify notation, write this as

$$\alpha_{ij}U_{i-1,j+1} + \beta_{ij}U_{i,j+1} - \alpha_{ij}U_{i+1,j+1} = V_i + c_5U_{ij} .$$

Hence, for $i = 1,2,\ldots,n-1$, we have $(n-1)$ equations in the $(n-1)$ unknowns

$$U_{1,j+1},\quad U_{2,j+1},\quad \ldots,\quad U_{n-1,j+1}.$$

These equations, together with the data, can be written in matrix form,

$$\text{(3.9)}\qquad FU = R ,$$

where

$$F = \begin{bmatrix} \beta_{1j} & -\alpha_{1j} & 0 & 0 & \cdots & 0 & 0 \\ \alpha_{2j} & \beta_{2j} & -\alpha_{2j} & 0 & \cdots & 0 & 0 \\ 0 & \alpha_{3j} & \beta_{3j} & -\alpha_{3j} & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & 0 & \cdots & \alpha_{n-1,j} & \beta_{n-1,j} \end{bmatrix},$$

$$U = \begin{bmatrix} U_{1,j+1} \\ U_{2,j+1} \\ \cdot \\ \cdot \\ U_{n-1,j+1} \end{bmatrix},$$

$$R = \begin{bmatrix} V_1 + c_5 U_{1j} - \alpha_{1j} A_{j+1} \\ V_2 + c_5 U_{2j} \\ \cdot \\ \cdot \\ V_{n-2} + c_5 U_{n-2,j} \\ V_{n-1} + c_5 U_{n-1,j} + \alpha_{n-1,j} B_{j+1} \end{bmatrix}.$$

The procedure is to let

$$c_5 U_{ij} = 0$$

and solve the matrix equation, (3.9), successively for

$$j = 1,2,\ldots,m-1 .$$

This gives the first approximate solution. The correction term, c_5U_{ij}, is then evaluated, included in the vector R and the corrected equation is solved, as before, for

$$j = 1,\ldots,m-1 .$$

This procedure is repeated until successive solutions agree to the required accuracy. An additional condition, usually applied when P.D. systems are to be solved by difference methods, is to require that differences (of the unknown function) of some order be less than a prescribed constant. This serves to determine the mesh sizes h_x and h_y.

§3.6.3. <u>The state of the art:</u>

Though the discussion of the preceding two sections was rather restricted in scope, it did serve to illustrate the manner in which P.D. systems are tackled numerically. We see that the problem is more difficult than a step from one to two dimensions would indicate and it is clear that the state of the art is nowhere as advanced as is the case in O.D.E 's. The analysis of truncation and round-off errors and stability is extremely difficult, particularly for quasi and non-linear systems.

The advent of digital computers has made the solution of two dimensional systems common and three dimensional systems possible, but computers have not cured all our ills. A great deal of time must be spent analysing and preparing the problem for the machine. Programs for solving P.D. systems are generally very complicated, involving

either a great deal of internal decision-making or input and output to permit human intervention. The great strength of digital computers is the ability to perform arithmetic operations rapidly and it is only with the greatest of difficulty that a computer can be programmed to even approximate human decision-making ability. In order to best make use of human intelligence as a machine saving device, considerable communication between the machine and its operator is required and as yet the speed of input/output devices lags far behind the operating speed of computers.

Writing the difference equation in matrix form, as we have done in §3.6.2, serves to emphasize the main drawback to decreasing the mesh size to ensure accuracy. The number of equations to be solved varies as $1/h_x h_y$ and the time taken to solve the systems of linear equations required to obtain the complete solution increases as the cube of the number of equations, that is, as $(1/h_x h_y)^3$. Furthermore, the more equations there are to be solved the greater the number of computations which must be performed; this increases round-off error and hence more significant figures must be carried, which serves to slow down computation. Finally, big matrices of numbers with many significant digits require large blocks of storage in the computer. If sufficient high speed storage (main memory) is not available, slow access storage (magnetic tape, for example) and more complicated methods of solving linear equations must be used. This again is costly in time.

Difference-correction techniques permit larger intervals, but they also suffer from certain drawbacks. In order to calculate differences near the data curve, values of the unknown function outside the region of solution must be known. These values are obtained either by

extrapolating the known solution or by solving the difference equation with already computed values as data. Both these procedures may be risky, especially if there is a possibility of crossing a characteristic curve. Since we are to eventually take differences of the exterior values, error would tend to be minimized, provided we know the solution to be continuous across the data curve. This is not always the case in initial value problems though it is usually true in elliptic systems. The evaluation of the difference-correction term requires a fairly complicated program and in addition memory requirements for the storage of the program and tables of differences are large, hence, auxiliary storage is usually necessary. All differences greater than some predetermined value must be included in the correction term. Since the magnitude of the difference depends on the mesh size, it may happen that the calculation of the correction term becomes tedious indeed if the mesh size is too large. A combination of decreasing-interval and difference-correction techniques appears to be the best approach; for example, we could require that differences greater than some given order are to be negligible and decrease the mesh size accordingly. In any case, runs of two or three days of continuous computing are not unusual.

Non-rectangular data curves make solving the difference equation extremely tedious since the curve will not pass through mesh points. If the mesh is adjusted to do so, then all grid rectangles are not the same size and closed difference methods cannot be applied so readily. One method of dealing with non-rectangular data curves is discussed in Chapter V. Sometimes it is possible to obtain a rectangular data curve by a change of dependent variables. In most cases the solution must be interpolated to the boundary.

It is clear that much basic work must be done to develop more efficient methods for solving P.D. systems and faster computers with increased memory capacity, decision-making ability and improved input/output facilities to implement these methods, before consideration of P.D.E 's can be brought to any degree of completion.

CHAPTER IV

SOLUTION OF THE EQUATION $\frac{\partial^2 U^5}{\partial x^2} = \frac{\partial U}{\partial t}$

§4.1. Introduction:

As an example of the application of difference methods to quasi-linear equations Richtmyer [14] considered the equation

$$(4.1) \qquad \frac{\partial^2 U^5}{\partial x^2} = \frac{\partial U}{\partial t}, \quad t \geq 0 .$$

By requiring U to be a running-wave solution of (4.1), that is,

$$U = U(x - vt) ,$$

Richtmyer obtained the following solution in implicit form:

$$\frac{5}{4}(U-U_o)^4 + \frac{20}{3} U_o(U-U_o)^3 + 15 U_o^2(U-U_o)^2 + 20 U_o^3(U-U_o)$$
$$+ 5 U_o^4 \ln(U-U_o) = v(vt-x+x_o) ,$$

where U_o, x_o and v are constants. Numerical values were assigned to these constants, the equation was solved for U by Newton's method and compared with values obtained by the numerical solution of (4.1). In general, the two solutions agreed closely.

A more direct solution to (4.1) is given in this chapter and Appendix I. It was hoped that a comparison of analytic and numerical solutions could be made here, but the task is difficult and time consuming and complete numerical results are not yet available. Some partial results are recorded briefly at the end of this chapter.

In this next section a brief discussion of the numerical

properties of parabolic systems will be given with particular reference to (4.1) and the heat equation

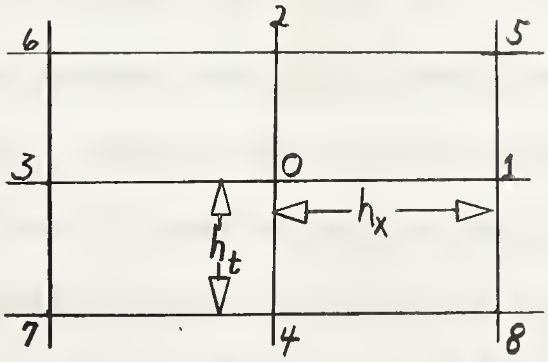
$$(4.2) \qquad a^2 \frac{\partial^2 U}{\partial x^2} = \frac{\partial U}{\partial t} \quad .$$

§4.2. The Parabolic Difference Equation:

§4.2.1. Implicit and explicit forms of the difference equation:

In approximating any P.D.E. by a difference equation, we have to decide which of the formulae of §2.2 to use. By considering the linear heat flow equation, (4.2), we will see that, in parabolic systems at least, this choice is crucial. As physical intuition might tell us, it is the essential difference between time-like and space-like co-ordinates which is the determining factor.

Established practice suggests that we replace the second order derivative by the central difference formula, (2.16). The time derivative, $\frac{\partial U}{\partial t}$, can be replaced by one of three formulae: forward (2.13), backward (2.14), or central (2.17); each of which results in a slightly different difference equation. To illustrate these three equations consider a typical mesh point in the region of solution and label it 0; surrounding mesh points are labelled 1, 2, 3, ... according to the figure below.

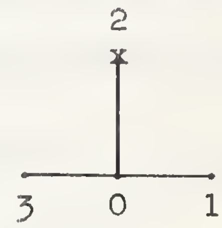


The positive x-direction is to the right and the positive t-direction is vertical. Let the value of U at the mesh point i by U_i. The mesh points involved in the three difference equations which follow appear in diagrams to the right of the equations. Lines joining points

indicate differences involving these points. Dots, (·), represent known points and crosses, (x), are unknown points. Only the first terms of the respective difference formulae are used.

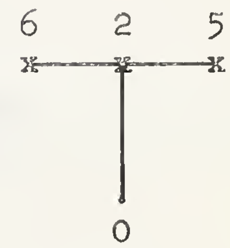
Using (2.13) and (2.16) at 0 gives

(4.3) $$h_t\, a^2(U_1 + U_3 - 2U_o) = h_x^2(U_2 - U_o) .$$



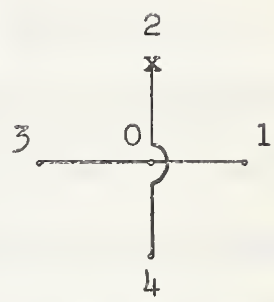
From (2.14) and (2.16) applied at 2 we have

(4.4) $$h_t\, a^2(U_5 + U_6 - 2U_2) = h_x^2(U_2 - U_o) .$$



Finally, if (2.17) and (2.16) are applied at 0, we obtain

(4.5) $$h_t\, a^2(U_1 + U_3 - 2U_o) = \frac{h_x^2}{2}(U_2 - U_4) .$$



Suppose that we are given the values of U along $t = 0$ and that we are carrying the solution forward in the direction of increasing t, one row at a time. Equation (4.5) is not suitable in this case since it requires two rows of data in order to get started. Equation (4.3) can be solved explicitly for U_2 provided U_3, U_o and U_1 are known. Assuming values at all the points ...,3,0,1,... are known we can use (4.3) repeatedly to obtain the next row, ...,6,2,5,... . Because the values on one row are given explicitly point by point in terms of values

on the preceding row, this formula is called explicit. If (4.4) is applied along the row ...,3,0,1,... we obtain a set of simultaneous algebraic equations which can be solved to give the solution along the row ...,6,2,5,... . Equation (4.4) is called implicit because, although the unknown values cannot be obtained directly from the equation, it does imply a set of equations from which the solution may be obtained.

It is possible to generalize (4.3) and (4.4) by taking their weighted average as follows: Let θ be a constant such that

$$0 \leq \theta \leq 1 .$$

Multiply (4.4) by θ, (4.3) by $(1-\theta)$ and add to get

$$(4.6) \qquad h_x^2(U_2-U_0) = h_t \; a^2[\theta\delta_x^2 U_2+(1-\theta)\delta_x^2 U_0] .$$

§4.2.2. Stability in partial difference systems:

Let us digress for a moment and consider the problem of stability in partial difference equations. In §2.4.3 we defined instability in an ordinary difference equation as the growth of parasitic solutions which eventually dominated the true solution, and we found that these parasitic solutions were introduced by an initial inaccuracy in the data. The properties of instability in partial and ordinary systems are quite similar. Boundary value problems (ordinary and partial) are generally stable, while initial value problems are subject to instability. Instability in partial systems generally manifests itself by a rapid oscillation of the numerical solution as well as an increasing departure from the analytic solution as one proceeds forward in the time direction. Reduction of the interval in the x-direction (h_x) without a corresponding reduction of the time interval (h_t) only magnifies the error. It has been observed that a drastic

reduction of the time interval reduces the effects of instability and that the system is stable for some values of h_x^2 / h_t and instable for others.

Quantitative investigations of instability in simple initial systems have been carried out [1,15,16,17]. Hildebrand [15] defines instability as the tendency for the effect of an initial numerical inaccuracy (such as round-off error) to increase unboundedly as the solution is carried forward in the t-direction.

§4.2.3. Stability in parabolic systems:

Among the most complete of stability analyses and of particular interest to us are the results of Courant, Friedrichs and Lewy [16] relating to the parabolic equation (4.2) with data

$$U(x,o) = \varphi(x) ,$$

$$U(o,t) = U(\pi,t) = 0$$

and a^2 a constant. It was found that the difference equation (4.6) is stable if

$$2\, a^2 \frac{h_t}{h_x^2} \leq \frac{1}{1-2\theta} \;, \quad 0 \leq \theta < \frac{1}{2} \;,$$

and with no restriction if

$$\frac{1}{2} \leq \theta \leq 1 \, .$$

The argument used to obtain this result is tedious and will not be given here.

Richtmyer [14] argues that the same conditions hold for the more general parabolic equation

$$\frac{\partial U}{\partial t} = a^2 \frac{\partial^2 U}{\partial x^2} + b \frac{\partial U}{\partial x} + c\,U ,$$

where a^2, b and c are constants, if the former is weakened to

$$2\,a^2 \frac{h_t}{h_x^2} < \frac{1}{1-2\theta} , \quad 0 \leq \theta < \tfrac{1}{2} .$$

By writing (4.1) in the form

$$\frac{\partial U}{\partial t} = 5\,U^4 \frac{\partial^2 U}{\partial x^2} + 20\,U^3 \frac{\partial U}{\partial x}$$

one can draw a comparison between a^2 and $5\,U^4$ and argue that the difference equation

$$(4.7) \qquad h_x^2\,(U_2 - U_0) = h_t[\theta\delta_x^2(U_2^5) + (1-\theta)\delta_x^2(U_0^5)]$$

is stable if

$$10\,U^4 \frac{h_t}{h_x^2} < \frac{1}{1-2\theta} , \quad 0 \leq \theta < \tfrac{1}{2} ,$$

with no restriction for $\frac{1}{2} \leq \theta \leq 1$. That this is indeed the case has been verified by numerical work [14].

From the form of (4.7) it is clear that, for other than $\theta = 0$, we must solve non-linear algebraic equations. However, if $\theta = 0$ the stability condition is

$$h_t < \frac{h_x^2}{10U^4}$$

and, unless $U << 1$, h_t must be extremely small if we are to avoid a very coarse x interval.

Hence, in the numerical work below we will employ the implicit scheme ($\theta=1$) and linearize the difference equation.

§4.3. Analytic Properties of the Equation $\frac{\partial^2 U^5}{\partial x^2} = \frac{\partial U}{\partial t}$:

§4.3.1. Analytic solution:

Writing

$$U(x,t) = X(x)\, T(t)$$

and substituting in (4.1) we obtain

$$T^5 \frac{d^2X^5}{dx^2} = X \frac{dT}{dt} \quad .$$

Hence,

$$\frac{1}{X} \frac{d^2X^5}{dx^2} = \frac{1}{T^5} \frac{dT}{dt} = -\lambda \ ,$$

where λ is a constant.

Consider first the equation for T,

$$\frac{1}{T^5} \frac{dT}{dt} = -\lambda \ .$$

This integrates to

$$-4\, T^{-4} = -\lambda\, t - c \ ,$$

where c is a constant.

Thus

$$T = \pm[4(c + \lambda t)]^{-4} \ .$$

If we associate U with temperature, we may assume that X and T are positive. We will require both X and T to be real and bounded. This means that both λ and c must be positive. Thus the t-dependent part of U is

$$T = [4(c + \lambda t)]^{-4} , \tag{4.8}$$

where c and λ are positive constants.

The equation for X is

$$\frac{d^2X^5}{dx^2} = - \lambda X ,$$

where X is positive. Let

$$Y = X^5 \geq 0 .$$

Then

$$\frac{d^2Y}{dx^2} = - \lambda Y^{1/5} .$$

Multiplying both sides of the above by $\frac{dY}{dx}$ and noting that

$$\frac{d^2Y}{dx^2} \frac{dY}{dx} = \frac{d}{dx} \left[\frac{1}{2} \left(\frac{dY}{dx}\right)^2\right]$$

we obtain

$$\frac{d}{dx} \left[\frac{1}{2} \left(\frac{dY}{dx}\right)^2\right] = - \lambda Y^{1/5} \frac{dY}{dx} .$$

After integrating with respect to x, we have

$$\left(\frac{dY}{dx}\right)^2 = \text{const.} - \frac{5}{3} \lambda Y^{6/5}$$

$$= \frac{5}{3} \lambda (a - Y^{6/5}) ,$$

where a is a constant. Since

$$T \geq 0 , \quad \lambda \geq 0 ,$$

then

$$a \geq 0 \quad .$$

Since Y is bounded then

$$\left(\frac{dY}{dx}\right)^2 \xrightarrow[x \to \infty]{} 0 \quad .$$

Hence,

$$a = \lim_{x \to \infty} Y^{6/5} \quad .$$

Let

$$Y = S^{5/2} \; .$$

Then

$$\frac{dY}{dx} = \frac{5}{2} S^{3/2} \frac{dS}{dx} \quad .$$

Hence, the last equation in Y can be written as

$$\frac{5}{2} S^{3/2} \frac{dS}{dx} = [\tfrac{5}{3} \lambda(a-S^3)]^{\frac{1}{2}} ,$$

or

$$\left(\frac{4\lambda}{15}\right)^{\frac{1}{2}} = \frac{S^{3/2}}{(a-S^3)^{\frac{1}{2}}} \frac{dS}{dx}$$

$$= \frac{S^2}{(aS-S^4)^{\frac{1}{2}}} \frac{dS}{dx} \quad .$$

If we integrate with respect to x we obtain

$$\left(\frac{4\lambda}{15}\right)^{\frac{1}{2}} x + b = \int_{S_1}^{S} \frac{S^2 \, dS}{(aS-S^4)^{\frac{1}{2}}} ,$$

where b and S_1 are constants. By letting

$$a = 8\eta^3 , \quad S = \eta\, u , \quad u_1 = S_1/\eta ,$$

we can write this last equation in the following form:

$$\left(\frac{4\lambda}{15}\right)^{\frac{1}{2}} x + b = \eta \int_{u_1}^{u} \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}} , \tag{4.9}$$

where

$$u = \frac{[X(x)]^2}{\eta} .$$

It can be shown that the integral

$$I = \int \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}}$$

reduces to a sum of elliptic integrals and the details of this reduction appear in Appendix I. It was originally intended that values of the analytic solution be compared with the numerical solution in order to illustrate the methods developed in this thesis. This would involve numerical evaluation of the integral in (4.9) or its equivalent elliptic integral sum. From a practical point of view the former is preferable since tables of elliptic integrals of the third kind are incomplete and interpolation in tables (where available) would be necessary in any case.

§4.3.2. Discussion of the analytic solution:

For the purposes of this section let us write equations (4.8) and (4.9) in a slightly different form as follows:

$$T(t) = (d + \mu t)^{-1/4} ,$$

$$\left(\frac{\mu}{15}\right)^{\frac{1}{2}} x + b = \eta \int_{u_1}^{X^2/\eta} \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}} , \tag{4.10}$$

where d, b, η, u_1 and μ are constants and

$$U(x,t) = X(x)\,T(t)$$

is the solution to the P.D.E. (4.1).

As pointed out in Chapter III, a P.D. problem is meaningful (in a practical sense) only if auxiliary conditions are imposed on the solution. Usually we must turn to the physics of the problem for appropriate conditions for it is not wise to apply arbitrary data since this often results in a system which is not well-posed. A physical analogue to (4.1) has not been discovered. In any case, the implicit nature of the second of (4.10) makes the application of auxiliary conditions difficult. Therefore, let us reverse the normal procedure and examine the solution so as to discover what data is appropriate to the equation.

We have at our disposal the constants d, b, η, u_1 and μ. The origin in x and t is determined by d and b, for

$$T(o) = d^{-1/4} \quad ,$$

$$b = \eta \int_{u_1}^{X^2(o)/\eta} \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}} \quad .$$

It is reasonable to take

$$T(o) = 1 \quad ,$$

that is,

$$d = 1 \ .$$

Since u_1 is as yet undetermined, we can let

$$b = 0 .$$

The rate at which the solution decays as t increases is determined by μ. It is convenient to set $\mu = 15$. This makes T decay fairly rapidly, but a full investigation would require a range of values.

The initial shape of the solution, $U(x,o)$, is, in the main, fixed by η. Again, a number of values should be tried, but $\eta = 1$ is a reasonable value to investigate.

Our solution now has the form

$$T = (1 + 15t)^{-1/4} ,$$

(4.10a)

$$x = \int_{u_1}^{X^2} \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}} .$$

We have required that X be a real function of x, hence, we must have

$$8u - u^4 \geq 0 .$$

Now

$$8u-u^4 = u(2-u)(u^2+2u+4) ,$$

but

$$u^2+2u+4 = [(u+1)^2+3]$$

which is always positive. Hence, we must have

$$0 \leq u \leq 2 .$$

It will now be shown that X is a periodic function of x. Let

$$w = u + iv$$

be a complex variable and consider the integral

$$I = \int_C \frac{w^2 dw}{(8w-w^4)^{\frac{1}{2}}} ,$$

where C is a "dumbbell" shaped contour composed of the circles, C_1 and C_2, about the branch points $w = 2$ and $w = 0$ joined by the lines AB and CD as shown in the figure below.

Suppose that the circles have radii ϵ and that the contour is so chosen that all other singularities are excluded. We may write

$$I = \int_{AB} \frac{w^2 dw}{(8w-w^4)^{\frac{1}{2}}} + \int_{CD} \frac{w^2 dw}{(8w-w^4)^{\frac{1}{2}}} + \int_{C_1} \frac{w^2 dw}{(8w-w^4)^{\frac{1}{2}}} + \int_{C_2} \frac{w^2 dw}{(8w-w^4)^{\frac{1}{2}}} .$$

On C_1 we write

$$2 - w = \epsilon\, e^{i\varphi} .$$

Then

$$\int_{C_1} = - i\, \epsilon^{\frac{1}{2}} \int_{-\pi}^{\pi} \frac{(2-\epsilon e^{i\varphi})^{3/2} e^{i\varphi/2}\, d\varphi}{[(2-\epsilon e^{i\varphi})^2+2(2-\epsilon e^{i\varphi})+4]^{\frac{1}{2}}}$$

and it is clear that this integral vanishes when $\epsilon \rightarrow 0$. Note that the function

$$\frac{w^2}{[w(w^2+2w+4)]^{\frac{1}{2}}}$$

is unchanged in sign as w passes around C_1 .

Now

$$(2-w)^{\frac{1}{2}} = \epsilon^{\frac{1}{2}} e^{i\varphi/2} .$$

When w is at A, $\varphi = -\pi$ and

$$(2-w)^{\frac{1}{2}} = -i .$$

When w is at C, $\varphi = \pi$ and

$$(2-w)^{\frac{1}{2}} = i .$$

Hence, $(2-w)^{\frac{1}{2}}$ changes sign as w goes around C_1. We can apply similar arguments for w on the circle C_2.

Thus, if we let

$$v \to 0, \quad \epsilon \to 0 ,$$

$$\int_{AB} \frac{w^2 dw}{(8w-w^4)^{\frac{1}{2}}} \to \int_0^2 \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}} ,$$

$$\int_{CD} \frac{w^2 dw}{(8w-w^4)^{\frac{1}{2}}} \to \int_0^2 \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}}$$

and

$$I = 2\int_0^2 \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}} .$$

Hence, as u goes through the range 0 to 2 to 0, x increases by

$$2 \int_0^2 \frac{u^2 du}{(8u-u^4)^{\frac{1}{2}}} .$$

Therefore, u is a periodic function of x with period

$$2\int_0^2 \frac{u^2du}{(8u-u^4)^{\frac{1}{2}}} \quad .$$

Since

$$X(x) = [u(x)]^{\frac{1}{2}} \quad ,$$

then X is a periodic function of x .

The point of the preceding analysis is this: just as we can investigate the function

$$y = \sin x$$

by considering the integral

$$x = \int_0^y \frac{dt}{(1-t^2)^{\frac{1}{2}}}$$

for

$$-1 \leq y \leq 1 \;,$$

$$0 \leq x \leq 2\pi = 2\int_{-1}^1 \frac{dt}{(1-t^2)^{\frac{1}{2}}} \quad ,$$

we can discover the properties of X(x) by considering

$$x = \int_0^{X^2} \frac{u^2du}{(8u-u^4)^{\frac{1}{2}}}$$

for

$$0 \leq X \leq 2 \quad ,$$

$$0 \leq x \leq 2\int_0^2 \frac{u^2du}{(8u-u^4)^{\frac{1}{2}}} \quad .$$

We can complete the problem by evaluating the above integrals numerically.

§4.4. A Numerical Procedure for Solving the Equation $\dfrac{\partial^2 U^5}{\partial x^2} = \dfrac{\partial U}{\partial t}$:

§4.4.1. Initial conditions:

In view of §4.3, a reasonable data curve to choose would be

$$t = 0, \quad x = x_1, \quad x = x_2 ,$$

where

$$x_2 > x_1 .$$

The initial conditions are

$$U(x,o) = A(x) ,$$

$$U(x_1,t) = B(t) ,$$

$$U(x_2,t) = C(t)$$

and the region of solution is

$$t \geq 0 , \quad x_1 \leq x \leq x_2 .$$

§4.4.2. The difference equation:

To avoid difficulties arising from instability let us use an implicit scheme ($\theta = 1$ in equation (4.7)). The difference equation is

$$k\, \delta_x^2\, U_{ij}^5 - \nabla_t\, U_{ij} = c_{ij} ,$$

where

$$U(x_1 + ih_x, jh_t) = U_{ij} ,$$

$$k = h_t/h_x^2 ,$$

$$c_{ij} = k\,(\delta_x^4 - \delta_x^6 + \ldots)\,U_{ij}^5 + (\tfrac{1}{2}\nabla_t^2 + \tfrac{1}{3}\nabla_t^3 + \ldots)U_{ij}\ ,$$

$$i = 1, 2, \ldots, n,$$

$$j = 1, 2, \ldots, m,$$

$$h_x = \frac{x_2 - x_1}{n+1}\ ,\quad h_t = \frac{1}{m+1}$$

and $\quad U_{io} = A_i,\ \ U_{oj} = B_j,\ \ U_{n+1,j} = C_j$.

We must now linearize the difference equation. This can be done in several ways. First, let us write the equation explicitly in terms of the unknown quantities,

$$U_{i-1,j},\ \ U_{ij},\ \ U_{i+1,j}$$

as follows:

$$k\,U_{i-1,j}^5 - 2k\,U_{ij}^5 - U_{ij} + k\,U_{i+1,j}^5 = -U_{i,j-1} + c_{ij}\ .$$

If we let α_i be an approximation to U_{ij}^4 in terms of known quantities, we have

$$k\alpha_{i-1}\,U_{i-1,j} + \beta_i\,U_{ij} + k\alpha_{i+1}\,U_{i+1,j} = -U_{i,j-1} + c_{ij}\ ,$$

where

$$\beta_i = -(2k\,\alpha_i + 1)\ .$$

The difference equation in matrix form is

(4.11) $\quad FZ = G\ ,$

where

$$F = \begin{bmatrix} \beta_1 & k\alpha_2 & 0 & 0 & . & . & . & 0 & 0 \\ k\alpha_1 & \beta_2 & k\alpha_3 & 0 & . & . & . & 0 & 0 \\ 0 & k\alpha_2 & \beta_3 & k\alpha_4 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & & . \\ 0 & 0 & 0 & 0 & . & . & . & k\alpha_{n-1} & \beta_n \end{bmatrix},$$

$$Z = \begin{bmatrix} U_{1j} \\ U_{2j} \\ . \\ . \\ . \\ U_{nj} \end{bmatrix}, \qquad G = \begin{bmatrix} c_{1j} - U_{1,j-1} - kB_j^4 \\ c_{2j} - U_{2,j-1} \\ . \\ . \\ . \\ c_{nj} - U_{n,j-1} - kC_j^4 \end{bmatrix}.$$

Note that α_o and α_{n+1} are known exactly, that is,

$$\alpha_o = U_{oj}^4 = B_j^4 ,$$

$$\alpha_{n+1} = U_{n+1,j}^4 = C_j^4 .$$

One linearization, based on the assumption that U is a slowly varying function of y, is to let

$$\alpha_i = U_{i,j-1}^4 .$$

We then compute the vector Z thereby obtaining a first approximation, ${}^{(1)}U_{ij}$. Next we let

$$\alpha_i = {}^{(1)}U_{ij}$$

and solve again to get a second and (we hope) improved approximation, ${}^{(2)}U_{ij}$. With

$$\alpha_i = {}^{(2)}U_{ij}$$

we solve (4.11) again, repeating the procedure until successive solutions agree to the required accuracy. Initially we must let $c_{ij} = 0$, but difference-corrections can be applied at any stage after the first.

Since the explicit difference equation is linear we can easily calculate ${}^{(1)}U_{ij}$. If we then let

$$\alpha_i = {}^{(1)}U_{ij} ,$$

we may be able to improve this first estimate by iterating as before. Although the explicit equation would probably be instable, the error should not be too large for only one time step.

Richtmyer [14] linearizes as follows: using Taylor's theorem,

$$U_{ij}^5 = U_{i,j-1}^5 + h_t \left(\frac{\partial U^5}{\partial t} \right)_{i,j-1} + \cdots$$

$$= U_{i,j-1}^5 + 5U_{i,j-1}^4 \, h_t \left(\frac{\partial U}{\partial t} \right)_{i,j-1} + \cdots ,$$

but

$$h_t \left(\frac{\partial U}{\partial t} \right)_{i,j-1} \approx \Delta_t U_{i,j-1}$$

$$\approx U_{i,j} - U_{i,j-1} .$$

Hence,

$$U_{ij}^5 \approx U_{i,j-1}^5 + 5U_{i,j-1}^4 (U_{i,j} - U_{i,j-1})$$

$$\approx (5U_{i,j-1}^4) \, U_{i,j} - 4U_{i,j-1}^5 .$$

§4.4.3. Some numerical results:

An algorithm was written to solve the following system:

$$\frac{\partial^2 U^5}{\partial x^2} = \frac{\partial U}{\partial t} ,$$

(4.12) $U(0,t) = U(1,t) = 0$,

$U(x,o) = \sin \pi x$,

$0 \leq x \leq 1$, $t \geq 0$.

The differential system was reduced to the difference system (4.11) and the first linearization scheme of §4.4.2 was used. The first approximation to the first row solution, ${}^{(1)}U_{i1}$, and three successive approximations, ${}^{(2)}U_{i1}$, ${}^{(3)}U_{i1}$, ${}^{(4)}U_{i1}$, are given in Table II. The mesh size is

$$h_x = h_t = 1/10 .$$

The solution is symmetric about $x = 0.5$ and only half of the solution is recorded.

TABLE II

Approximate solution of the system (4.12) along the line $t = 1/10$.

x	${}^{(1)}U(x,1/10)$	${}^{(2)}U(x,1/10)$	${}^{(3)}U(x,1/10)$	${}^{(4)}U(x,1/10)$
0	0.0000	0.0000	0.0000	0.0000
0.1	1.1271	0.0317	0.3165	0.6881
0.2	0.8576	0.1377	2.0933	0.0269
0.3	0.5169	0.7398	0.5076	0.5729
0.4	0.3805	1.1569	0.1525	1.3261
0.5	0.3441	1.1598	0.1738	0.9961

The successive solutions show no tendency to converge. Decreasing the mesh size effects some change, but mesh sizes of the order of 1/100 would be necessary for significant improvement. Clearly the linearization procedure used is too crude. It is likely that the choice of initial data is inappropriate, introducing a rapidly varying dependence on t into the solution.

Consideration of this problem is far from complete, but the foregoing has served to point out the inadvisability of considering the differential equation as a mathematical abstraction with arbitrary initial conditions. Lacking a physical basis for a meaningful problem suitable data will have to be obtained numerically from the analytic solution using various values for the constants; only then will it be possible to undertake the numerical solution.

CHAPTER V

ELLIPTIC SYSTEMS

§5.1. Introduction:

The wide application of closed, difference methods to the solution of boundary value problems in ordinary differential equations [5] would lead one to suspect that similar methods are commonly used in the solution of elliptic systems. This is not the case. To date, matrix methods have been applied only to a limited class of linear equations with very simple boundaries, whereas relaxation procedures [3] and sweeping techniques [20] (recent improvements of relaxation) have received more attention and have been successfully adapted to many types of equations and boundaries. Nevertheless, matrix methods offer a conciseness of formulation and an invariance of procedure which are lacking in sweeping techniques.

In this chapter some of the matrix-difference methods for solving linear elliptic P.D. systems are discussed. At first only very simple equations are considered, but generalizations are made wherever possible. We assume that the boundary curve is a rectangle with sides parallel to the axes and that the solution is specified on the boundary. Non-rectangular boundaries and boundary conditions involving the normal derivative of the solution are considered separately. One section of this chapter is devoted to a discussion of singularities in boundary conditions, but otherwise it is assumed that the solution and data are continuous and non-singular. In most cases, concepts are illustrated by particular examples. In what follows we will use results from the theory of matrices [6,18]; no proofs for these theorems will be given.

§5.1.1. Conventions:

To avoid unnecessary repetition in later work let us establish the following conventions:

The region of solution of the equation is the rectangle

$$a_o \leq x \leq a_1 \quad , \quad b_o \leq y \leq b_1 \quad .$$

A grid or mesh is set up in this region consisting of the points

$$(x_i \; , \; y_j) \; ,$$

where

$$x_i = a_o + ih_x \; ,$$

$$y_j = b_o + jh_y \; ,$$

$$i = 0,1,\ldots,n+1,$$

$$j = 0,1,\ldots,m+1,$$

$$h_x = \frac{a_1-a_o}{n+1} \quad , \quad h_y = \frac{b_1-b_o}{m+1} \quad .$$

Interior points are those points, (x_i,y_j), for which

$$i = 1,2,\ldots,n,$$

$$j = 1,2,\ldots,m$$

and boundary points are

$$(x_o,y_j), \; (x_{n+1},y_j), \; (x_i,y_o), \; (x_i,y_{m+1}) \; ,$$

$$i = 0,1,\ldots,n+1 \; ,$$

$$j = 0,1,\ldots,m+1 \; .$$

The value of the solution at the point (x_i, y_j) is denoted by z_{ij}. The matrix of the solution at interior points is denoted by Z or $\{z_{ij}\}$. We will also use the letter Z to represent the continuous solution whenever a P.D. system is written down. It should be clear from the context whether Z stands for a matrix of discrete values or the continuous solution.

Other conventions will be given as we need them.

§5.2. Matrix Representations of Some Simple Elliptic Systems:

Among the most common of elliptic systems are those involving Poisson's equation. The method used to reduce this equation to a matrix system can easily be extended to handle more general types of elliptic P.D.E's and, hence, will serve as a model for later work.

Let us write Poisson's equation in the form

$$(5.1) \qquad -\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right) = f(x,y).$$

The reason for the minus sign will be apparent later. Multiply (5.1) by $(h_x^2 h_y^2)$ and use the central difference formula (2.16) to obtain the difference equation

$$(5.2) \qquad -(h_y^2\delta_x^2 z_{ij} + h_x^2\delta_y^2 z_{ij}) = h_x^2 h_y^2 f_{ij} + (h_y^2 c_x + h_x^2 c_y) z_{ij} ,$$

where c_x and c_y are difference-correction terms.

It is established custom to have

$$h_x = h_y = h$$

unless it is known in advance or by a trial computation that the solution varies more rapidly in one direction than in the other.

Let

$$(c_x + c_y)z_{ij} = c_{ij} ,$$

$$h^2 f_{ij} + c_{ij} = g_{ij} .$$

Then (5.2) can be written

$$(5.3) \quad \begin{aligned} &-(\delta_x^2 z_{ij} + \delta_y^2 z_{ij}) = g_{ij} , \\ &i = 1,2,\ldots,n , \\ &j = 1,2,\ldots,m . \end{aligned}$$

A double array of quantities such as (5.3) can be most conveniently represented by a matrix. Let us consider the special case $n = 3$, $m = 4$ to see if a compact matrix representation of (5.3) can be found. In matrix form, therefore, we have

$$\begin{bmatrix} -z_{01}+2z_{11}-z_{21} & -z_{02}+2z_{12}-z_{22} & -z_{03}+2z_{13}-z_{23} & -z_{04}+2z_{14}-z_{24} \\ -z_{11}+2z_{21}-z_{31} & -z_{12}+2z_{22}-z_{32} & -z_{13}+2z_{23}-z_{33} & -z_{14}+2z_{24}-z_{34} \\ -z_{21}+2z_{31}-z_{41} & -z_{22}+2z_{32}-z_{42} & -z_{23}+2z_{33}-z_{43} & -z_{24}+2z_{34}-z_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} -z_{10}+2z_{11}-z_{12} & -z_{11}+2z_{12}-z_{13} & -z_{12}+2z_{13}-z_{14} & -z_{13}+2z_{14}-z_{15} \\ -z_{20}+2z_{21}-z_{22} & -z_{21}+2z_{22}-z_{23} & -z_{22}+2z_{23}-z_{24} & -z_{23}+2z_{24}-z_{25} \\ -z_{30}+2z_{31}-z_{32} & -z_{31}+2z_{32}-z_{33} & -z_{32}+2z_{33}-z_{34} & -z_{33}+2z_{34}-z_{35} \end{bmatrix}$$

$$= \begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \end{bmatrix} .$$

Now, the two matrices on the left-hand side of the equation above contain the given boundary data,

z_{oj}, z_{4j}, z_{io}, z_{i5},

$i = 1,2,3,4,$

$j = 1,2,3,4,5.$

Let us retain only unknown quantities on the left side of the equation by adding the matrix

$$\begin{bmatrix} z_{01}+z_{10} & z_{02} & z_{03} & z_{04}+z_{15} \\ z_{20} & 0 & 0 & z_{25} \\ z_{41}+z_{30} & z_{42} & z_{43} & z_{44}+z_{35} \end{bmatrix}$$

to both sides. The matrix equation is now

$$\mathcal{Z}_1 + \mathcal{Z}_2 = F ,$$

where

$$\mathcal{Z}_1 = \begin{bmatrix} 2z_{11}-z_{21} & 2z_{12}-z_{22} & 2z_{13}-z_{23} & 2z_{14}-z_{24} \\ -z_{11}+2z_{21}-z_{31} & -z_{12}+2z_{22}-z_{32} & -z_{13}+2z_{23}-z_{33} & -z_{14}+2z_{24}-z_{34} \\ -z_{21}+2z_{31} & -z_{22}+2z_{32} & -z_{23}+2z_{33} & -z_{24}+2z_{34} \end{bmatrix},$$

$$\mathcal{Z}_2 = \begin{bmatrix} 2z_{11}-z_{12} & -z_{11}+2z_{12}-z_{13} & -z_{12}+2z_{13}-z_{14} & -z_{13}+2z_{14} \\ 2z_{21}-z_{22} & -z_{21}+2z_{22}-z_{23} & -z_{22}+2z_{23}-z_{24} & -z_{23}+2z_{24} \\ 2z_{31}-z_{32} & -z_{31}+2z_{32}-z_{33} & -z_{32}+2z_{33}-z_{34} & -z_{33}+2z_{34} \end{bmatrix},$$

$$F = \begin{bmatrix} g_{11}+z_{01}+z_{10} & g_{12}+z_{02} & g_{13}+z_{03} & g_{14}+z_{04}+z_{15} \\ g_{21}+z_{20} & g_{22} & g_{23} & g_{24}+z_{25} \\ g_{31}+z_{41}+z_{30} & g_{32}+z_{42} & g_{33}+z_{43} & g_{34}+z_{44}+z_{35} \end{bmatrix}.$$

The matrix $\mathcal{Z}_1$ can be written as the product

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{21} & z_{22} & z_{23} & z_{24} \\ z_{31} & z_{32} & z_{33} & z_{34} \end{bmatrix}$$

and $\mathcal{Z}_2$ can be written as the product

$$\begin{bmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{21} & z_{22} & z_{23} & z_{24} \\ z_{31} & z_{32} & z_{33} & z_{34} \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} .$$

Following Bickley and McNamee [19], let

$$D_3^2 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} ,$$

$$D_4^2 = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} ,$$

where the subscript denotes the order of the matrix. Thus, we can write the matrix equation as

$$D_3^2 \; Z + Z \, D_4^2 = F .$$

Let D_n^2 be the triple diagonal matrix of order $(n x n)$ whose diagonal elements are 2 and whose sub-and super-diagonal elements are -1. Then, by the same procedure as employed in the example above, we can write (5.3) in matrix form

$$(5.4) \qquad D_n^2 \, Z + Z \, D_m^2 = F ,$$

where Z and F are matrices of order $(n \times m)$. F consists of the elements

$$\{h^2 f_{ij} + c_{ij}\} ,$$

with the quantities

$$z_{oj},\ z_{n+1,j},\ z_{io},\ z_{i,m+1}$$

added to the first row, last row, first column and last column respectively, where

$$i = 1,2,\ldots,n,$$
$$j = 1,2,\ldots,m.$$

It should be pointed out that equation (5.4) represents the complete difference system (equation with boundary values). We also note that D_n^2 is a positive-definite matrix. It was to assure the positive-definiteness of D_n^2 that we chose the minus sign in (5.1).

In precisely the same way we can represent the equations

$$\text{(5.5)} \qquad -\left(c(x)\,\frac{\partial^2 z}{\partial x^2} + e(y)\,\frac{\partial^2 z}{\partial y^2}\right) = f(x,y)$$

and

$$\text{(5.6)} \qquad -\left(\frac{\partial^2[c(x)Z]}{\partial x^2} + \frac{\partial^2[e(y)Z]}{\partial y^2}\right) = f(x,y)$$

by

$$\text{(5.4a)} \qquad CD_n^2\,Z + Z\,D_m^2\,E = F$$

and

$$\text{(5.4b)} \qquad D_n^2\,CZ + ZED_m^2 = F$$

respectively, where C and E are diagonal matrices of order $(n\times n)$ and $(m\times m)$ respectively.

§5.3. Solution of the Matrix-Difference Equation:

Equations (5.4), (5.4a), (5.4b) can be written in the form

$$AZ + ZB = F \tag{5.7}$$

where A is of order $(n\times n)$, B is of order $(m\times m)$ and Z and F are of order $(n\times m)$. A, B and F are known.

Three methods of obtaining the solution of (5.7) were given by Bickley and McNamee [19]. The first of these involves the eigenquantities* of both A and B and is called the irrational method. The second, termed semi-rational, requires the eigenquantities of A or B but not both. The third, named rational, does not, in theory, require eigenquantities at all.

If a matrix of order $(k\times k)$ has k linearly independent eigenvectors, it is defined to be non-defective. In the irrational solution both A and B must be non-defective, while the semi-rational method requires that at least one of them be non-defective. In the third method no restriction is placed on A or B. The reason for the names irrational, semi-rational and rational will become apparent later.

§5.3.1. The irrational solution:

Suppose that A and B are non-defective and that A has the n eigenvalues λ_i, the n eigenrows ρ_i and the n eigen-

---

* Eigenquantities-eigenvalues and eigenvectors.

columns Γ_i while B has the m eigenvalues ν_j, the m eigenrows σ_j and the m eigencolumns s_j.

The eigenquantities of A obey the following relations:

$$\rho_i A = \lambda_i \rho_i ,$$

$$A \Gamma_i = \lambda_i \Gamma_i , \tag{5.8}$$

$$\rho_i \Gamma_j = \delta_{ij} .$$

The eigenquantities of B obey the relations

$$\sigma_j B = \nu_j \sigma_j ,$$

$$B s_j = \nu_j s_j , \tag{5.9}$$

$$\sigma_i s_j = \delta_{ij} .$$

Let u_{ij} be scalars, where

$$i = 1,2,\ldots,n ,$$

$$j = 1,2,\ldots,m .$$

The theory of matrices tells us that an arbitrary matrix can be represented as a linear combination of products of the eigenvectors of non-defective matrices. In particular

$$Z = \sum_{i=1}^{n} \sum_{j=1}^{m} u_{ij} \Gamma_i \sigma_j .$$

If we substitute this in (5.7) we have

$$A \sum_{i,j} u_{ij} \Gamma_i \sigma_j + \sum_{i,j} u_{ij} \Gamma_i \sigma_j B = F .$$

In view of (5.8) and (5.9) this becomes

$$\sum_{i,j} (\lambda_i + \nu_j)\, u_{ij}\, \Gamma_i\, \sigma_j = F \; .$$

Now, if we pre-multiply this relation by ρ_k, post-multiply by s_ℓ and use the last of (5.8) and (5.9), we obtain

$$(\lambda_k + \nu_\ell)\, u_{k\ell} = \rho_k\, F\, s_\ell \; .$$

Hence, provided

$$(\lambda_k + \nu_\ell) \neq 0 \; ,$$

we can find the scalars u_{ij} from

$$u_{ij} = \frac{\rho_i\, F s_j}{(\lambda_i + \nu_j)} \; ,$$

$$i = 1,2,\ldots,n \; ,$$

$$j = 1,2,\ldots,m \; .$$

This method requires that we find the eigenquantities of both A and B, which is no mean task. So, except for a few special cases (in particular D_n^2), this method is to be avoided in practical computation.

§5.3.2. <u>The semi-rational solution</u>:

Let

$$U_j \; , \quad j = 1,2,\ldots,m$$

be columns of order (nx1). Suppose that B is non-defective, but place no such restriction on A. We can expand the solution matrix, Z, in the form

$$Z = \sum_{j=1}^{m} U_j \sigma_j \ .$$

Equation (5.7) becomes

$$A \sum_j U_j \sigma_j + \sum_j U_j \sigma_j B = F.$$

After application of the first of (5.9) we have

$$A \sum_j U_j \sigma_j + \sum_j \nu_j U_j \sigma_j = F \ .$$

Post-multiply by s_i and use the last of (5.9) to get

$$A U_i + \nu_i U_i = F s_i \ .$$

Therefore, if we define I to be the matrix with unit diagonal and zero off-diagonal, U_i can be obtained by solving the equation

$$(A + \nu_i I) U_i = F s_i \ ,$$

provided

$$|A + \nu_i I| \neq 0 \ .$$

This is equivalent to the condition

$$(\lambda_j + \nu_i) \neq 0 \ .$$

It is clear that a similar relation can be found in terms of the eigenquantities of the (non-defective) matrix A.

The semi-rational solution is of more practical value than the irrational solution even though n sets of m linear equations in m unknowns must be solved, for obtaining eigenquantities accurately is exceedingly tedious.

§5.3.3. The rational solution:

Before proceeding with the rational solution of (5.7), a few results from matrix theory should be stated. We defined the eigen-quantities of a matrix B such that they satisfy the relations

$$\sigma_i B = \nu_i \sigma_i ,$$

$$B s_i = \nu_i s_i .$$

These equations can be written as

$$\sigma_i (B - \nu_i I) = 0 ,$$

$$(B - \nu_i I) s_i = 0$$

and hence, the determinant

$$|B - \nu_i I| = 0 .$$

Expanding the determinant gives us a polynomial in ν which is called the characteristic polynomial of the matrix B and is written

$$\nu^m - p_1 \nu^{m-1} + \ldots + (-1)^m p_m = 0$$

(5.10) or

$$(\nu-\nu_1)(\nu-\nu_2) \ldots (\nu-\nu_m) = 0.$$

The Cayley-Hamilton theorem states that a matrix satisfies its own characteristic equation, that is,

$$B^m - p_1 IB^{m-1} + \ldots + (-1)^m p_m I = 0$$

(5.10a) or

$$(B-\nu_1 I)(B-\nu_2 I) \ldots (B-\nu_m I) = 0 ,$$

where 0 is the zero matrix.

The rational solution proceeds as follows: let ν_1 be any constant. Add $\nu_1 IZ$ to both sides of (5.7) to obtain

$$(5.11) \qquad (A+\nu_1)Z = F + Z(\nu_1 - B) ,$$

where $(A+\nu_1)$ implies $(A+\nu_1 I)$. Let ν_2 be a constant and multiply (5.11) by $(A+\nu_2)$ to get

$$(5.12) \qquad (A+\nu_2)(A+\nu_1)Z = AF + \nu_2 F + \nu_2 Z(\nu_1 - B) + \nu_1 AZ - AZB .$$

But from (5.7)

$$AZ = F - ZB .$$

Hence,

$$AZB = FB - ZB^2$$

and

$$\nu_1 AZ = \nu_1 F - \nu_1 ZB .$$

Substituting these last relations in (5.12) we have

$$\begin{aligned}(A+\nu_2)(A+\nu_1)Z &= (\nu_1+\nu_2)F + (AF-FB) + Z[\nu_2(\nu_1-B) - \nu_1 B + B^2] \\ &= (\nu_1+\nu_2)F + (AF-FB) + Z(\nu_1-B)(\nu_2-B) .\end{aligned}$$

Multiplying this by $(A + \nu_3)$ and proceeding as before we obtain

$$\begin{aligned}(5.13) \qquad (A+\nu_3)(A+\nu_2)(A+\nu_1)Z &= (\nu_1\nu_2+\nu_2\nu_3+\nu_3\nu_1)F + (\nu_1+\nu_2+\nu_3)(AF-FB) \\ &\quad +(A^2F-AFB+FB^2) + Z(\nu_1-B)(\nu_2-B)(\nu_3-B) .\end{aligned}$$

We can continue in this way with constants $\nu_4, \nu_5, \ldots, \nu_m$, but to see more clearly where this analysis is leading consider the special case $m = 3$. From (5.10a) we have

$$(\nu_1-B)(\nu_2-B)(\nu_3-B) = 0 \quad .$$

From (5.10) we see that the coefficient of F in (5.13) is p_2 and that the coefficient of $(AF-FB)$ is p_1. Now, let

$$F(k) = A^kF - A^{k-1}FB + A^{k-2}FB^2 - \cdots + (-1)^k FB^k \quad .$$

We can write (5.13) in the form

$$(A+\nu_3)(A+\nu_2)(A+\nu_1)Z = p_2 F(0) + p_1 F(1) + F(2) \quad .$$

For general m, it follows that

$$(A+\nu_m)(A+\nu_{m-1})\cdots(A+\nu_1)Z = \sum_{k=o}^{m-1} p_{m-k-1} F(k) \quad .$$

If we let

$$(A+\nu_m)(A+\nu_{m-1})\cdots(A+\nu_1) = C \quad ,$$

the solution for Z is

$$Z = C^{-1} \sum_{k=o}^{m-1} p_{m-k-1} F(k) \quad .$$

Clearly, we could have obtained a similar result in terms of the matrix A and in general we will use the matrix of lower order.

Note that the matrix C can also be given in terms of the coefficients of the characteristic polynomial and so a knowledge of the eigenvalues of A or B is not necessary. The coefficients of the characteristic equation can (theoretically) be calculated by purely rational operations, hence this method is called rational. The eigenquantities of a matrix cannot be obtained by rational means. There

are, however, no good numerical methods for finding the coefficients of the characteristic equation which involve appreciably less labor than finding the eigenvalues and the distinction between rational and irrational has little practical significance. The second method is called semi-rational merely to distinguish it from the first.

§5.4. Matrix Representation of More General Elliptic Equations:

In §5.2 and §5.3 we showed how a restricted class of linear elliptic systems could be represented in terms of matrix equations. The purpose of this section is to reduce equations of the form

$$-\left(a(x,y)\frac{\partial^2 z}{\partial x^2} + b(x,y)\frac{\partial^2 z}{\partial y^2}\right) = f(x,y) \tag{5.14}$$

to matrix equations and to put forward one method for solving them, provided a and b are slowly varying functions of x and y.

§5.4.1. Matrix equation for (5.14):

The difference equation for (5.14) is

$$-(a_{ij}\,\delta_x^2\, z_{ij} + b_{ij}\,\delta_y^2\, z_{ij}) = g_{ij}\ , \tag{5.15}$$

where

$$g_{ij} = h^2 f_{ij} + \text{correction terms.}$$

We define the matrix dot product, $C\cdot E = L$, as follows: let

$$C = \{c_{ij}\},\quad E = \{e_{ij}\},\quad L = \{\ell_{ij}\}\ .$$

Then

$$\ell_{ij} = c_{ij}\, e_{ij}\ ,$$

that is, corresponding elements of C and E are multiplied together. The result (L) is a matrix.

If we write the complete set of difference equations (5.15) in matrix form as we did in §5.2, we obtain the matrix equation

$$(5.16)\qquad A \cdot (D_n^2 Z) + (Z D_m^2) \cdot B = F ,$$

where

$$A = \{a_{ij}\} , \quad B = \{b_{ij}\}$$

and

$$F = \{h^2 f_{ij} + \text{correction terms}\}$$

incremented in the first row, last row, first column and last column by

$$a_{1j}\, z_{oj},\ a_{nj}\, z_{n+1,j},\ b_{i1}\, z_{io} \ \text{ and } \ b_{im}\, z_{i,m+1}$$

respectively.

§5.4.2. <u>Solution of the matrix equation (5.16)</u>:

Equation (5.16) is to be solved for the (nxm) matrix Z, but the nature of the dot product does not permit direct application of the methods of §5.3. Let us examine the dot product $A \cdot Y$ for the special case $n = 2$, $m = 3$.

Suppose, first of all, that A has the special form

$$A = \begin{bmatrix} a_{11} & a_{11} & a_{11} \\ a_{22} & a_{22} & a_{22} \end{bmatrix} .$$

Then

$$A \cdot Y = \begin{bmatrix} a_{11}y_{11} & a_{11}y_{12} & a_{11}y_{13} \\ a_{22}y_{21} & a_{22}y_{22} & a_{22}y_{23} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \end{bmatrix} ,$$

that is,

$$A \cdot Y = A_D\, Y ,$$

where A_D is a diagonal matrix of order (2x2). Now, if A is of the more general form

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} ,$$

we can write

$$A = A_1 + A_2 ,$$

where

$$A_1 = \begin{bmatrix} \alpha_1 & \alpha_1 & \alpha_1 \\ \alpha_2 & \alpha_2 & \alpha_2 \end{bmatrix} ,$$

$$A_2 = \begin{bmatrix} a_{11}-\alpha_1 & a_{12}-\alpha_1 & a_{13}-\alpha_1 \\ a_{21}-\alpha_2 & a_{22}-\alpha_2 & a_{23}-\alpha_2 \end{bmatrix} .$$

Thus

$$A \cdot Y = A_1 \cdot Y + A_2 \cdot Y$$

$$= A_{1D}\, Y + A_2 \cdot Y .$$

Similarly, the dot product $Y \cdot B$ can be written

$$Y \cdot B = YB_{1D} + Y \cdot B_2 ,$$

where

$$B_{1D} = \begin{bmatrix} \beta_1 & 0 & 0 \\ 0 & \beta_2 & 0 \\ 0 & 0 & \beta_3 \end{bmatrix} ,$$

$$B_2 = \begin{bmatrix} b_{11}-\beta_1 & b_{12}-\beta_2 & b_{13}-\beta_3 \\ b_{22}-\beta_1 & b_{22}-\beta_2 & b_{23}-\beta_3 \end{bmatrix} .$$

Hence in the case $n = 2$, $m = 3$ equation (5.16) can be written as

$$A_{1D}D_2^2Z + ZD_3^2B_{1D} = F - [A_2\cdot(D_2^2Z) + (ZD_3^2)\cdot B_2] .$$

This suggests the general form to which (5.15) can be reduced, namely:

$$(5.17) \qquad A_{1D}D_n^2Z + ZD_m^2B_{1D} = F - [A_2\cdot(D_n^2Z) + (ZD_m^2)\cdot B_2] ,$$

where A_{1D} is a diagonal matrix of order (nxn) with diagonal elements

$$\alpha_1, \alpha_2, \ldots, \alpha_n ,$$

B_{1D} is diagonal of order (mxm) with elements

$$\beta_1, \beta_2, \ldots, \beta_m ,$$

A_2 is of order (nxm) whose ith row is the ith row of A decreased by α_i and B_2 is (nxm) whose jth column is the jth column of B decreased by β_j .

If $a(x,y)$ and $b(x,y)$ are slowly varying functions, we may be able to choose the α_i and β_j in such a way that the term

$$[A_2\cdot(D_n^2Z) + (ZD_m^2)\cdot B_2]$$

can be considered to be a small perturbation. In this case we can ignore the perturbing term and solve

$$A_{1D}D_n^2 Z + ZD_m^2 B_{1D} = F$$

by the methods of §5.3. Having obtained Z we can evaluate the perturbing term, include it in (5.17) and solve for a better estimate of Z. Continuing in this manner we may expect that the procedure will, under certain conditions, converge to the required solution.

It must be pointed out that this method is as yet untried and the problem of convergence and the selection of the α_i and β_j has not been investigated. If the elements of the perturbing matrix are small compared to the elements of F we may reasonably expect the procedure to converge fairly rapidly.

§5.5. Solution of Elliptic Systems by the Big Matrix:

Probably the most straightforward method for solving linear elliptic equations, provided that the solution is specified on the boundary, is to write the complete set of difference equations in the form

$$\mathcal{B} Z = F ,$$

where Z is a column vector of the solution, F is a column vector of the boundary values and the inhomogeneous term of the P.D.E. and is the matrix of difference coefficients. Following Bickley and McNamee [19], we will call $\mathcal{B}$ the big matrix.

§5.5.1. The big-matrix representation of the Laplace and Poisson equations:

To illustrate the use of the big matrix let us consider Laplace's equation

$$-\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right) = 0 .$$

The corresponding difference equation is

$$-(\delta_x^2 z_{ij} + \delta_y^2 z_{ij}) = c_{ij} ,$$

where c_{ij} is the correction term and z_{oj}, $z_{n+1,j}$, z_{io}, $z_{i,m+1}$ are known.

Let us consider the case of twelve interior points corresponding to

$$i = 1,2,3,$$

$$j = 1,2,3,4.$$

Write the difference equation in the following way:

$$-z_{i,j+1} - z_{i-1,j} + 4z_{ij} - z_{i+1,j} - z_{i,j+1} = c_{ij} ,$$

$$i = 1,2,3; \quad j = 1,2,3,4.$$

This notation implies twelve equations for the values of the solution at the twelve interior points. These equations can be written in matrix notation

$$\mathcal{B} Z = F ,$$

where Z is a column of unknowns*,

$$Z^t = (z_{11}, z_{21}, z_{31}, z_{12}, z_{22}, z_{32}, z_{13}, z_{23}, z_{33}, z_{14}, z_{24}, z_{34}) ,$$

F is a column of right-hand sides,

---

* Superscript t indicates the transpose of the matrix.

$$F^t = (c_{11}+z_{10}+z_{01}, c_{21}+z_{20}, c_{31}+z_{30}+z_{41}, c_{12}+z_{02}, c_{22}, c_{32}+z_{42},$$

$$c_{13}+z_{03}, c_{13}+z_{03}, c_{23}, c_{33}+z_{43}, c_{14}+z_{04}+z_{15}, c_{24}+z_{25},$$

$$c_{34}+z_{44}+z_{35}) ,$$

and $\mathcal{B}$ is the matrix of coefficients,

$$\mathcal{B} = \left[\begin{array}{ccc|ccc|ccc|ccc}
4 & -1 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
-1 & 4 & -1 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & -1 & 4 & \cdot & \cdot & -1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
-1 & \cdot & \cdot & 4 & -1 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & -1 & \cdot & -1 & 4 & -1 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & -1 & \cdot & -1 & 4 & \cdot & \cdot & -1 & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & -1 & \cdot & \cdot & 4 & -1 & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & -1 & \cdot & -1 & 4 & -1 & \cdot & -1 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & -1 & 4 & \cdot & \cdot & -1 \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & \cdot & 4 & -1 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & -1 & 4 & -1 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & -1 & 4
\end{array}\right]$$

Let

$$P_3 = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix} , \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} , \quad 0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} .$$

The big matrix can be partitioned into submatrices and written

$$\mathcal{B} = \begin{bmatrix} P_3 & -I & 0 & 0 \\ -I & P_3 & -I & 0 \\ 0 & -I & P_3 & -I \\ 0 & 0 & -I & P_3 \end{bmatrix} .$$

By following the same procedure for general n and m we obtain

$$\mathcal{B} Z = F ,$$

where

$$\mathcal{B} = \begin{bmatrix} P_n & -I & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ -I & P_n & -I & 0 & \cdots & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 0 & -I & P_n & -I \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -I & P_n \end{bmatrix},$$

$$Z^t = (z_{11}, z_{21}, \ldots, z_{n1}, z_{12}, z_{22}, \ldots, z_{n2}, \ldots, z_{1m}, z_{2m}, \ldots, z_{nm}) ,$$

$$F^t = (c_{11}+z_{10}+z_{01}, c_{21}+z_{20}, \ldots, c_{n1}+z_{n0}+z_{n+1,1}, c_{12}+z_{02}, c_{22}, c_{32},$$
$$\ldots, c_{n-1,2}, c_{n2}+z_{n+1,2}, c_{13}+z_{03}, \ldots, c_{nm}+z_{n+1,m}+z_{n,m+1}) .$$

$\mathcal{B}$ is of order (nmxnm) and Z and F are of length nm. The matrices P_n, I and 0 are of order (nxn). I is the unit matrix, 0 is the zero matrix and

$$P_n = \begin{bmatrix} 4 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 4 & -1 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & -1 & 4 \end{bmatrix} .$$

The big-matrix equation corresponding to Poisson's equation, (5.1), is

$$\mathcal{B} Z = F + G ,$$

where

$$G^t = h^2(f_{11}, f_{21}, \ldots, f_{n1}, f_{12}, \ldots, f_{nm}) .$$

§5.5.2. Generalizations and discussion:

Clearly, any linear elliptic P.D.E. can be written in the big-matrix form and, provided a suitable linearization can be made, quasi- and non-linear elliptic equations, too, are reducible to a big-matrix equation. Obviously the elements of the big matrix will depend upon the coefficients of the difference equation, making it difficult to write down the general form of $\mathcal{B}$. The procedure for setting up the matrix equation simply consists of writing down the difference equation at each of the nm points of the region of solution.

It is apparent that this method suffers from serious drawbacks which make it of dubious practical value. Even with the application of difference-corrections a required accuracy of four or more figures seldom makes consideration of fewer than one hundred interval points practical. This means that one hundred linear equations must be solved. Furthermore, the number of matrix elements to be stored exceeds the memory capacity of most computers and, to prevent loss of significance by round-off during computation, so many figures must be carried that computing time would be prohibitive. The simplicity of the big-matrix formulation masks its crudeness. It is a brutal statement of the problem, but it is the most general closed difference method and by its crudeness points out how little is known about solving P.D. systems.

§5.6. Differential Boundary Conditions:

Often the boundary conditions accompanying elliptic equations involve the normal derivative of the solution. Derivative conditions as such cannot be directly incorporated since we must have

values of the solution itself on the boundary. The procedure usually followed is exemplified in the following example involving Poisson's equation*:

$$-\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right) = f(x,y) ,$$

$$\frac{\partial z}{\partial n} = 0 \quad \text{on the boundary} \quad x = a_o,\ x = a_1,\ y = b_o,\ y = b_1 .$$

The difference equation is

$$(5.18) \qquad -(\delta_x^2 z_{ij} + \delta_y^2 z_{ij}) = h^2 f_{ij} + c_{ij} ,$$

where now

$$i = 0,1,2,\ldots,n+1 ,$$

$$j = 0,1,2,\ldots,m+1 ,$$

n and m as defined in §5.1.1, that is, the difference equation must be applied at points on the boundary and involves values of the solution outside the boundary. The boundary conditions are

$$\left(\frac{\partial z}{\partial x}\right)_{oj} = \left(\frac{\partial z}{\partial x}\right)_{n+1,j} = \left(\frac{\partial z}{\partial y}\right)_{io} = \left(\frac{\partial z}{\partial y}\right)_{i,m+1} = 0 .$$

Using (2.17),

$$h\,D = \mu\,\delta + c'$$

$$= \frac{1}{2}\,(E-E^{-1}) + c',$$

---

* It is easily possible to construct more complicated examples in which considerable effort must be expended in finding the difference equation at nodal points on or near the boundary, but this would only serve to obscure the main point of this section.

where c' is a correction term, the boundary conditions are

$$(5.19)\qquad \begin{aligned} z_{-1,j} &= z_{1j}+c'_{oj},\quad z_{n+2,j} = z_{n,j}+c'_{n+1,j}\ ,\\ z_{i,-1} &= z_{i1}+c''_{io},\quad z_{i,m+2} = z_{im}+c''_{i,m+1}\ ,\end{aligned}$$

where

$$c'_{oj} = 2c'_x z_{oj}\ ,\quad c'_{n+1,j} = -2c'_x z_{n+1,j}\ ,$$

$$c''_{io} = 2c'_y z_{io}\ ,\quad c''_{i,m+1} = -2c'_y z_{i,m+1}\ .$$

We can write the difference equations in matrix form in a manner similar to that of §5.2. Let us consider the special case $n=1$, $m=2$. The matrix equation is

$$\mathcal{Z}_1 + \mathcal{Z}_2 = G\ ,$$

where

$$\mathcal{Z}_1 = \begin{bmatrix} -z_{-1,0}+2z_{00}-z_{10} & -z_{-1,1}+2z_{01}-z_{11} & -z_{-1,2}+2z_{02}-z_{12} & -z_{-13}+2z_{03}-z_{13} \\ -z_{00}+2z_{10}-z_{20} & -z_{01}+2z_{11}-z_{21} & -z_{02}+2z_{12}-z_{22} & -z_{03}+2z_{13}-z_{23} \\ -z_{10}+2z_{20}-z_{30} & -z_{11}+2z_{21}-z_{31} & -z_{12}+2z_{22}+z_{32} & -z_{13}+2z_{23}-z_{33} \end{bmatrix},$$

$$\mathcal{Z}_2 = \begin{bmatrix} -z_{0,-1}+2z_{00}-z_{01} & -z_{00}+2z_{01}-z_{02} & -z_{01}+2z_{02}-z_{03} & -z_{02}+2z_{03}-z_{04} \\ -z_{1,-1}+2z_{10}-z_{11} & -z_{10}+2z_{11}-z_{22} & -z_{11}+2z_{12}-z_{13} & -z_{12}+2z_{13}-z_{14} \\ -z_{2,-1}+2z_{20}-z_{21} & -z_{20}+2z_{21}-z_{22} & -z_{21}+2z_{22}-z_{23} & -z_{22}+2z_{23}-z_{24} \end{bmatrix},$$

$$G = \{h^2 f_{ij} + c_{ij}\}\ ,\quad i = 0,1,2;\quad j = 0,1,2,3.$$

The values $z_{-1,j}$, $z_{3,j}$, $z_{i,-1}$, $z_{i,4}$ are replaced by the expressions (5.19) to give

$$\mathcal{Z}_1 = \begin{bmatrix} 2z_{00}-2z_{10}-c'_{00} & 2z_{01}-2z_{11}-c'_{01} & 2z_{02}-2z_{12}-c'_{02} & 2z_{03}-2z_{13}-c'_{03} \\ -z_{00}+2z_{10}-z_{20} & -z_{01}+2z_{11}-z_{21} & -z_{02}+2z_{12}-z_{22} & -z_{03}+2z_{13}-z_{23} \\ -2z_{10}+2z_{20}-c'_{20} & -2z_{11}+2z_{21}-c'_{21} & -z_{12}+2z_{22}-c'_{22} & -2z_{13}+2z_{23}-c'_{23} \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -2 & 0 \\ -1 & 2 & -1 \\ 0 & -2 & 2 \end{bmatrix} \begin{bmatrix} z_{00} & z_{01} & z_{02} & z_{03} \\ z_{10} & z_{11} & z_{12} & z_{13} \\ z_{20} & z_{21} & z_{22} & z_{23} \end{bmatrix} - \begin{bmatrix} c'_{00} & c'_{01} & c'_{02} & c'_{03} \\ 0 & 0 & 0 & 0 \\ c'_{20} & c'_{21} & c'_{22} & c'_{23} \end{bmatrix}$$

and

$$\mathcal{Z}_2 = \begin{bmatrix} 2z_{00}-2z_{01}-c''_{00} & -z_{00}+2z_{01}-z_{02} & -z_{01}+2z_{02}-z_{03} & -2z_{02}+2z_{03}-c''_{03} \\ 2z_{10}-2z_{11}-c''_{10} & -z_{10}+2z_{11}-z_{12} & -z_{11}+2z_{12}-z_{13} & -2z_{12}+2z_{13}-c''_{13} \\ 2z_{20}-2z_{21}-c''_{20} & -z_{20}+2z_{21}-z_{22} & -z_{21}+2z_{22}-z_{23} & -2z_{22}+2z_{23}-c''_{23} \end{bmatrix}$$

$$= \begin{bmatrix} z_{00} & z_{01} & z_{02} & z_{03} \\ z_{10} & z_{11} & z_{12} & z_{13} \\ z_{20} & z_{21} & z_{22} & z_{23} \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 0 \\ -2 & 2 & -1 & 0 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

$$- \begin{bmatrix} c''_{00} & 0 & 0 & c''_{03} \\ c''_{10} & 0 & 0 & c''_{13} \\ c''_{20} & 0 & 0 & c''_{23} \end{bmatrix} .$$

Let D_k^3 be the matrix of order $(k\mathrm{x}k)$ defined by

$$D_k^3 = \begin{bmatrix} 2 & -2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -2 & 2 \end{bmatrix} .$$

Then the matrix equation is

$$D_3^3 Z + Z(D_4^3)^t = F ,$$

where

$$F = \begin{bmatrix} g_{00}+c'_{00}+c''_{00} & g_{01}+c'_{01} & g_{02}+c'_{02} & g_{03}+c'_{03}+c''_{03} \\ g_{10}+c''_{10} & g_{11} & g_{12} & g_{13}+c''_{13} \\ g_{20}+c'_{20}+c''_{20} & g_{21}+c'_{21} & g_{22}+c'_{22} & g_{23}+c'_{23}+c''_{23} \end{bmatrix} ,$$

$$g_{ij} = h^2 f_{ij} + c_{ij} .$$

Clearly the matrix equation for general n and m is

$$(5.20) \qquad D_{n+2}^3 Z + Z(D_{m+2}^3) = F ,$$

where

$$Z = \{z_{ij}\}$$

and

$$F = \{h^2 f_{ij} + c_{ij}\}$$

incremented in the first row by c'_{oj}, in the last row by $c'_{n+1,j}$, in the first column by c''_{io} and in the last column by $c''_{i,m+1}$, for

$$i = 0,1,2,\ldots,n+1 ,$$

$$j = 0,1,2,\ldots,m+1 .$$

Equation (5.20) can be solved by the methods of §5.3. A first approximation is obtained for all correction terms equal to zero. The solution is then extrapolated to give values outside the boundary, correction terms are evaluated and included in the right-hand side of (5.20) and the equation is solved again. This procedure is continued until successive solutions agree to the required accuracy.

§5.7. The Treatment of Singular Points:

Up to now we have assumed that the boundary and initial data are non-singular. Three types of singularity were discussed by Woods [21] in relation to Poisson's equation, namely, discontinuities in the derivative of the solution (type I), logarithmic singularities in the solution (type II) and simple discontinuities in the solution (type III). These are probably the most common types of singularity and the method of dealing with them will exemplify how we may handle the problem in general.

Suppose we are given Poisson's equation,

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = f(x,y) \quad ,$$

with data containing singularities of the types defined above. Since the Laplacian operator is linear we can always make the substitution

$$U(x,y) = \varphi(x,y) + \psi(x,y) \ .$$

The function ψ is chosen to be a solution of Laplace's equation and is such that it contains the singularities leaving φ non-singular. We then solve the problem

$$\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} = f(x,y)$$

with non-singular data in φ. The final solution, being the sum of φ and ψ, is then calculable except possibly, for a neighbourhood of the singularity.

In the case of Poisson's equation the choice of ψ is relatively simple since we know that the real and imaginary parts of a complex variable are solutions to Laplace's equation. Suppose the point (r_o,θ_o) is a singular point. If (r_o,θ_o) is a type I singular point, we choose

$$\psi = (r-r_o)^m (\alpha \sin m\,\theta + \beta \cos m\,\theta)^m ,$$

and, if m is not an integer, ψ has a singular derivative of some order at (r_o,θ_o).

For a type II singularity at (r_o,θ_o) we let

$$\psi = \mathrm{Re}\,[\log\,((r-r_o)e^{i\theta})^m] = m \log\,(r-r_o) .$$

Finally, a type III singularity determines

$$\psi = \mathrm{Im}\,[\log((r-r_o)\,e^{i\theta})^m] = m\,\theta .$$

A similar procedure is used for other linear equations except that a function ψ cannot always be found which both contains the singularities and satisfies the homogeneous equation, hence, the equation in φ differs from the equation in U in the inhomogeneous term. This is no disadvantage provided singularities are not introduced into the equation by this term.

As far as elliptic equations are concerned, type I and III singularities cause no trouble if closed methods of solution are used and need not be removed. Type II singularities must be taken care of in the manner described. Because of the marching procedure used in initial value problems all varieties of singularity cause an instability which is propagated in the direction of increasing time.

The following example will illustrate the technique discussed above:

Suppose we are required to solve

$$xU_{xx} - yU_{xy} + U_{yy} = 0 , \tag{5.21}$$

$$U(0,y) = |y| , \quad \left(\frac{\partial U}{\partial x}\right)_{x=0} = -1, \quad U(x,\pm 1) = 1 ,$$

$$-1 \leq x \leq 0 \quad , \quad -1 \leq y \leq 1 .$$

The equation is hyperbolic in the region of solution since

$$\Delta = \frac{y^2}{4} - x > 0 \quad .$$

Because

$$U(0,y) = |y| ,$$

there is a type I singularity at the origin. If we differentiate

$$U(x,\pm 1) = 1$$

with respect to x, we have

$$\left(\frac{\partial U}{\partial x}\right)_{\substack{x=0\\ y=\pm 1}} = 0 \quad ,$$

whereas the second initial condition gives us

$$\left(\frac{\partial U}{\partial x}\right)_{\substack{x=o \\ y=\pm 1}} = -1 .$$

Hence, there is a type I singularity at the points

$$x = 0, \quad y = \pm 1 .$$

Let

$$U = \varphi + \psi_1 + \psi_2 ,$$

where

$$\psi_1 = \frac{2}{\pi}\left\{ x \tan^{-1}\left(\frac{x}{1-y^2}\right) - \left(\frac{1-y^2}{2}\right) \log[x^2+(1-y^2)^2] \right\} ,$$

$$\psi_2 = |y| .$$

The system for φ is

$$x\varphi_{xx} - y\varphi_{xy} + \varphi_{yy} + \frac{2}{\pi}\left\{\log[x^2+(1-y^2)^2] + \frac{y^2(6y^2-3x-8) + x(1+2x) + 2}{x^2 + (1-y^2)^2}\right\} = 0,$$

$$\varphi(0,y) = \frac{2}{\pi}(1-y^2) \log (1-y^2) ,$$

$$\varphi(x,\pm 1) = -x ,$$

$$\left(\frac{\partial \varphi}{\partial x}\right)_{x=o} = -1 .$$

The numerical solution of this system can be undertaken with confidence using the method of §3.6.2. The final solution to (5.21),

$$U_{ij} = \varphi_{ij} + (\psi_1)_{ij} + (\psi_2)_{ij} ,$$

will hold everywhere in the given region except for a neighbourhood of the line $y = 0$.

§5.8. Non-rectangular Boundaries:

If the curve on which boundary data is given is not a rectangle with sides parallel to the axes,the methods discussed in the preceding sections of this chapter are useless. Curved boundaries do not present such a great problem to relaxation methods [3,20] since the mesh size can be varied to fit the boundary. If closed methods are used, the usual procedure is to use the given conditions to estimate data on a rectangular boundary. A crude solution can thereby be obtained and a better estimate of the rectangular conditions can be made using this solution. This process is repeated until sufficient accuracy is obtained.

The above technique, while proven by many successful solutions to be a valuable numerical procedure, lacks the concise statement and invariant form which distinguishes closed methods. Such procedures are not easily programmed for a digital computer since they depend, to a great extent, on the intuition and experience of the computor rather than a fixed set of rules. The estimation of the first set of conditions on the rectangular boundary is crucial and not easily undertaken by a casual computor.

In the following section a more straightforward and easily implemented method for handling non-rectangular boundaries is suggested.

§5.8.1. The method of linear combination of solutions:

When solving P.D. systems analytically we often represent the general solution as a sum of particular solutions to the equation. We might reasonably expect that the solution to a difference system

could be obtained by combining a finite number of solutions to the difference equation in a manner determined by the boundary conditions. The problem is to obtain solutions to the difference equation and we have seen that this is not possible in general without referring to boundary conditions. Then if we must solve a difference system, let us do so for simple boundary conditions on a simple boundary and combine a number of solutions obtained in this manner in such a way that they satisfy the given conditions at a number of points on the given boundary. We might then reasonably expect this combination to be a solution of the given difference system and an approximate solution to the differential system. The most suitable boundary is a rectangle with sides parallel to the axes and the simplest boundary conditions consist of values of the solution.

In particular, let us state the method as follows:

We are given the linear elliptic equation

(5.22) $L(Z) = 0$

and boundary conditions

(5.23) $Z = z_B$

on the boundary B given by $g(x,y) = 0$.

We choose a rectangular boundary R given by

$$x = a_0,\ x = a_1,\ y = b_0,\ y = b_1$$

which encloses B. In R we set up a mesh of nm interval points according to §5.1.1. The difference equation is

(5.24) $L'(z_{ij}) = 0$.

This equation is to be solved k times for k different boundary conditions on R. The choice of conditions is arbitrary, but one of the more obvious is to make the solution zero on the boundary except at one mesh point where it takes on the value unity. By moving this point to k different positions we obtain k different boundary conditions. The mesh lines intersect R at $2(m+n)$ points (corner points do not appear in the difference equation and so are not considered). Hence, in this case

$$k \leq 2(m + n) \quad .$$

The matrix difference equation can now be written down and solved by the methods of §5.3, §5.4 or §5.5. We obtain k solutions

$$(5.25) \qquad z^{(1)}, \; z^{(2)}, \; \ldots, \; z^{(k)} \; .$$

Let Z be the matrix of the final solution. We choose k points on B and call them

$$(5.26) \qquad B_1, \; B_2, \; \ldots, \; B_k \; .$$

From (5.23) we have the following known quantities:

$$z_{B_1} \; , \; z_{B_2} \; , \; \ldots, \; z_{B_k} \; .$$

Each of the solutions (5.25) can be interpolated to the choosen points (5.26) on B to give the k^2 quantities

$$\begin{array}{llll} z_{B_1}^{(1)}, & z_{B_2}^{(1)}, & \ldots, & z_{B_k}^{(1)} \; , \\ z_{B_1}^{(2)}, & z_{B_2}^{(2)}, & \ldots, & z_{B_k}^{(2)} \; , \\ \cdot & \cdot & \cdot & \cdot \\ z_{B_1}^{(k)}, & z_{B_2}^{(k)}, & \ldots, & z_{B_k}^{(k)} \; . \end{array}$$

We postulate constants

$$a_1, a_2, \ldots, a_k$$

such that

$$\begin{aligned}
a_1 z_{B_1}^{(1)} + a_2 z_{B_1}^{(2)} + \ldots + a_k z_{B_1}^{(k)} &= z_{B_1} , \\
a_1 z_{B_2}^{(1)} + a_2 z_{B_2}^{(2)} + \ldots + a_k z_{B_2}^{(k)} &= z_{B_2} , \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad & \cdot \\
a_1 z_{B_k}^{(1)} + a_2 z_{B_k}^{(2)} + \ldots + a_k z_{B_k}^{(k)} &= z_{B_k} .
\end{aligned}$$

These k linear equations in k unknowns can be solved for

$$a_1, a_2, \ldots, a_k .$$

The final solution, Z, is given by the following linear combination of the matrices (5.25):

$$Z = \sum_{i=1}^{k} a_i Z^{(i)} . \tag{5.27}$$

In a sense, this is equivalent to the spectral resolution of the matrix Z which we discussed in §5.3.1. We let

$$Z = \sum_{i=1}^{n} \sum_{j=1}^{m} u_{ij} \Gamma_i \sigma_j ,$$

where Γ_i and σ_j were known eigenvectors and the u_{ij} were chosen to make Z satisfy the difference equation and one set of boundary conditions. In a similar manner we can write (5.27) in the form

$$Z = \sum_{\ell=1}^{k} \sum_{i=1}^{n} \sum_{j=1}^{m} u_{ij\ell} \Gamma_i \sigma_j ,$$

where the $u_{ij\ell}$ are chosen to make Z satisfy the difference equation and $(k + 1)$ sets of boundary conditions.

Certain drawbacks to this method must be pointed out. Computing time has been increased considerably and storage for k additional (nxm) matrices is required. Some care must be exercised in the choice of the conditions on R so that we obtain independent solutions.

§5.8.2. Example:

Consider Laplace's equation in the quarter circle,

$$\frac{\partial^2 Z}{\partial x^2} + \frac{\partial^2 Z}{\partial y^2} = 0 \,, \tag{5.28}$$

$$\begin{aligned} &Z(x,0) = x^4, \\ &Z(0,y) = y^4, \\ &Z(x,y) = 1 \text{ on } x^2 + y^2 = 1, \\ &0 \le x \le 1, \quad 0 \le y \le 1, \quad x^2 + y^2 \le 1 \,. \end{aligned} \tag{5.28a}$$

The analytic solution is

$$Z(r,\theta) = r^4 \cos(4\theta) \tag{5.29}$$

in polar co-ordinates, or

$$Z(x,y) = (x^2+y^2)^2 \cos(4 \tan^{-1} \frac{y}{x})$$

in Cartesian co-ordinates. More difficult examples could be chosen but they would not demonstrate the numerical procedure any better. A simpler example to investigate would have been the system with solution

$$r^2 \cos (2\theta) \,,$$

but this is also the exact solution to the difference system and so would not be a fair test of the method.

Since the given boundary involves the circumference of the quarter circle, application of the method of §5.8.1 is appropriate. For the rectangular boundary, R, let us choose the square

$$x = 0, \quad x = 1, \quad y = 0, \quad y = 1.$$

Following the conventions of §5.1.1 we set up a mesh of $(n+2)^2$ boundary and interior points (x_i, y_j), where

$$x_i = ih, \quad y_j = jh,$$
$$i,j = 0,1,2,\dots,n+1,$$
$$h = \frac{1}{n+1} .$$

According to §5.8.1 we must choose k sets of boundary conditions on R. The matrix equation to be solved is (see §5.2)

$$D_n^2 Z^{(\ell)} + Z^{(\ell)} D_n^2 = F^{(\ell)} , \tag{5.30}$$

$$\ell = 1,2,\dots,k$$

which will give us k solutions. The boundary conditions on R are taken to be

$$\begin{aligned} z_{io}^{(\ell)} &= (ih)^4 , \\ z_{oj}^{(\ell)} &= (jh)^4 , \\ z_{i,n+1}^{(\ell)} &= \delta_{\ell i} , \\ z_{n+1,j}^{(\ell)} &= \delta_{\ell, j+n} , \end{aligned} \tag{5.30a}$$

where

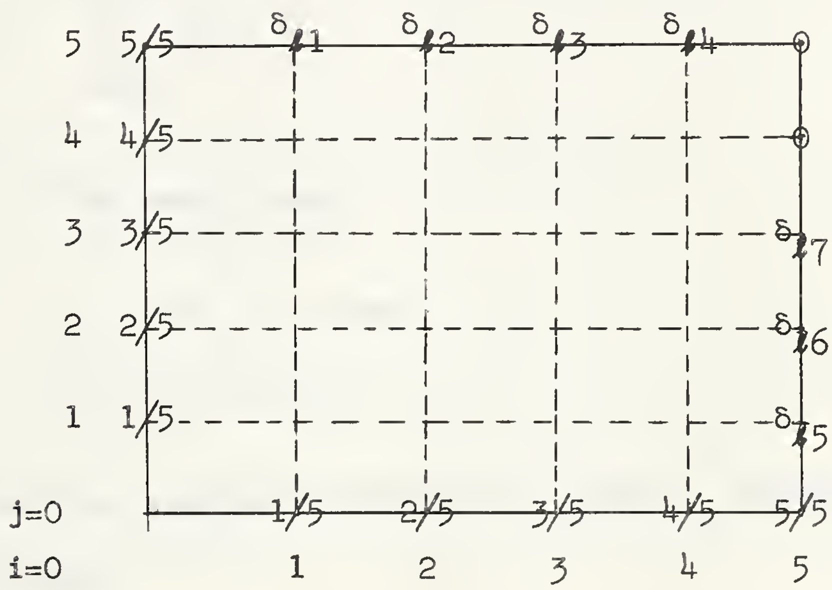
$\ell = 1, 2, \ldots, 2n-1$,

$\delta_{\ell m} = 0$ for $\ell \neq m$

and

$\delta_{mm} = 1$.

We have taken $k = 2n-1$. The figure below shows the region of solution for $n = 4$. Boundary values appear at their appropriate mesh points.



The first two conditions impose a restriction on the constants $a_1, a_2, \ldots, a_{2n-1}$ of §5.8.1. We see this as follows: the final solution, Z, is given by (5.27),

$$Z = \sum_{\ell=1}^{2n-1} a_\ell \, Z^{(\ell)} .$$

But

$$z_{io}^{(\ell)} = z_{io} = (ih)^4$$

and

$$z_{oj}^{(\ell)} = z_{oj} = (jh)^4 ,$$

that is, the final solution and all the sub-solutions have the same values along the axes. Hence,

$$z_{io} = \sum_{\ell=1}^{2n-1} a_\ell \, z_{io}^{(\ell)}$$

$$= z_{io} \sum_{\ell=1}^{2n-1} a_\ell \, .$$

Thus,

$$(5.31) \qquad \sum_{\ell=1}^{2n-1} a_\ell = 1 \, .$$

The mesh lines,

$$x_i = ih, \quad y_j = jh,$$

$$i,j = 1,2,\ldots,n,$$

intersect the quarter circle in $2n$ points. However, by choosing n such that

$$(n+1)^2 = s^2 + t^2 \, ,$$

where s and t are integers

$$1 \leq s, \quad t \leq n$$

we see that there are actually $(2n-2)$ points of intersection. Hence, with (5.31) we have $(2n-1)$ conditions on the $(2n-1)$ constants a_ℓ. Labelling these points of intersection

$$B_1, B_2, \ldots, B_{2n-2}$$

each of the $(2n-1)$ solutions, $z^{(\ell)}$, can be interpolated to each

of these points. Then, since the boundary values on $x^2+y^2 = 1$ are unity, we have that

$$\sum_{\ell=1}^{2n-1} a_\ell Z_{B_k}^{(\ell)} = 1 \ , \quad k = 1,2,\ldots,2n-2,$$

$$\sum_{\ell=1}^{2n-1} a_\ell = 1 \ .$$

These equations can be written in matrix form

$$(5.32) \qquad A = B$$

and solved to give the vector of constants

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{2n-1} \end{bmatrix} .$$

The final solution, (5.27), is then easily obtained.

The above procedure was programmed in Fortran for the I.B.M. 1620 computer. Since the eigenquantities are known (see Appendix II) equation (5.30) was solved by the irrational method of §5.3.1. As sufficient storage was not available in memory each sub-solution, $Z^{(\ell)}$, was punched out on cards as calculated. Each matrix was read back into memory when called for by the computer and interpolation to the quarter circle was performed. Four terms of a standard backward difference interpolation formula were used. Equation (5.32) was solved and the final solution formed. The

numerical solution so obtained was compared with the analytic solution (5.29). The numerical solution and the difference between numerical and analytic solution were punched on cards and listed on the 407 line printer. Difference-corrections were not applied. Running time, as a function of n, was found empirically to be

$$t = .004n^4 + .1n^3 \text{ minutes.}$$

Running time for $n = 9$ was about one and a half. The solution for $n = 9$, to four decimal places, appears in Table III. Values at every second mesh point are listed. The figures in brackets are the differences between the analytic and numerical solutions.

TABLE III

Approximate solution of the system (5.28), (5.28a).

Mesh spacing = 0.2. Every second value listed.

y \ x	0.0	0.2	0.4	0.6	0.8	1.0
1.0	1.0000 (00)	0.7614 (01)	0.0651 (04)			
0.8	0.4096 (00)	0.2578 (-02)	-0.1786 (-05)	-0.8436 (04)	-1.2321 (02)	
0.6	0.1296 (00)	0.0456 (-08)	-0.1890 (-13)	-0.5174 (-09)	-0.8435 (03)	
0.4	0.0256 (00)	-0.0098 (-13)	-0.1007 (-16)	-0.1890 (-13)	-0.1787 (-04)	0.0652 (03)
0.2	0.0016 (00)	-0.0053 (-10)	-0.0099 (-12)	0.0456 (-08)	0.2577 (-01)	0.7613 (02)
0.0	0.0000 (00)	0.0016 (00)	0.0256 (00)	0.1296 (00)	0.4096 (00)	1.0000 (00)

Considering the crudeness of the interval and the fact that no difference corrections were used, the obtained accuracy of two decimals is quite good. In order to improve significantly the accuracy of the solution, difference-corrections must be applied. Because of the fourth order dependence of time on mesh size, reduction of mesh spacing to improve accuracy would be prohibitively costly in time.

It is clear that this method for handling non-rectangular boundaries requires further investigation. Rates of convergence and the effect of the choice of matching points on the given boundary should be considered. It is likely, however, that this procedure can be used with about as much confidence as any other closed difference method. Boundary conditions involving derivatives of the solution can be similarly handled.

CHAPTER VI

CONCLUSION

In the preceding chapters an attempt has been made to describe how one goes about solving a second order partial differential system in two independent and one dependent variables. Much of what has been said can be applied to equations of higher order and more independent variables. In the way of a general method, all that can be said is that the differential system is changed to a difference system by replacing derivatives by terms from a difference series and the difference system is solved numerically. This procedure has been illustrated briefly in reference to hyperbolic and parabolic equations (Chapters III and IV) and in more detail with respect to elliptic systems (Chapter V). Examples have been given.

Except for a few special cases, the numerical solution cannot be justified by a priori arguments. We expect that trouble will manifest itself in some obvious way such as instability or non-convergence of successive approximations to the solution (Chapter IV).

Matrix methods for solving elliptic systems have been discussed in detail (Chapter V). Where applicable, these methods provide a straightforward and easily implemented algorithm for obtaining a solution.

On the whole, each partial differential system must be considered on its own merits which clearly indicates that our knowledge of partial differential equations - analytical and numerical - is far from complete.

BIBLIOGRAPHY

[1] G. E. Forsythe and W. R. Wasow: Finite Difference Methods for Partial Differential Equations. Wiley, (1960).

[2] W. G. Bickley: Difference and Associated Operators with Some Applications. Journal of Mathematics and Physics, Vol. XXVII, No. 3, (Oct. 1948).

[3] R. V. Southwell: Relaxation Methods in Theoretical Physics. Vol. 1 (1946), Vol. 2 (1956), Oxford Univ. Press.

[4] L. Fox: Some Improvements in the Use of Relaxation Methods for the Solution of Ordinary and Partial Differential Equations. Proc. Roy. Soc. A, 190, 31-59, (1947).

[5] L. Fox: The Numerical Solution of Two-Point Boundary Problems in Ordinary Differential Equations. Oxford, (1957).

[6] D. R. Hartree: Numerical Analysis. Clarendon Press, (1958).

[7] F. G. Tricomi: Differential Equations. Hafner, (1961).

[8] A. M. Turing: Rounding-off Errors in Matrix Processes. Quart. J. Mech. and Applied Math. Vol. 1, part 3, 287-308, (1948).

[9] J. H. Wilkinson: Rounding Errors in Algebraic Processes. Information Processing, UNESCO, 44-53, (1960).

[10] R. Courant and D. Hilbert: Methods of Mathematical Physics, Vol. II. Interscience, (1962).

[11] I. G. Petrovsky: Lectures on Partial Differential Equations. Interscience, (1961).

[12] J. Hadamard: Lectures on Cauchy's Problem in Linear Partial Differential Equations. Dover, (1952), first printed in 1923.

[13] Notes on Applied Science No. 16, Modern Computing Methods. H. M. Stationery Office, 2nd Edition, (1961).

[14] R. D. Richtmyer: Difference Methods for Initial Value Problems. Interscience, (1957).

[15] F. B. Hildebrand: Methods of Applied Mathematics. Prentice-Hall, (1952).

[16] R. Courant, K. O. Friedrichs and H. Lewy: Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann., vol. 100, p.32, (1928).

[17] John Todd: A Direct Approach to the Problem of Stability in the Numerical Solution of Partial Differential Equations. Communications on Pure and Applied Mathematics, Vol. IX, 597-612, (1956).

[18] M. Marcus: Basic Theorems in Matrix Theory. N.B.S. Applied Math. Series 57, (June, 1960).

[19] W. G. Bickley and J. McNamee: Matrix and Other Direct Methods for the Solution of Systems of Linear Difference Equations. Phil. Trans. Roy. Soc., Series A, No. 1005, Vol. 252, p.69, (Jan. 21,1960).

[20] H. B. Keller: On Some Iterative Methods for Solving Elliptic Difference Equations. Quart. Appl. Math. 16, 209-226, (1958).

[21] L. C. Woods: The Relaxation Treatment of Singular Points in Poisson's Equation. Quart. Journ. Mech and Applied Math., Vol. VI, pt.2, (1953).

[22] E. T. Whittaker and G. N. Watson: A Course of Modern Analysis. Cambridge University Press, 4th edition, (1940) first printed in 1902.

[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]

[illegible]
[illegible]
[illegible]

APPENDIX I

REDUCTION OF THE INTEGRAL $\int \frac{u^2\,du}{\sqrt{8u-u^4}}$ TO STANDARD ELLIPTIC INTEGRAL FORM

We have

$$I = \int \frac{u^2\,du}{\sqrt{8u-u^4}} \quad .$$

Now

$$8u - u^4 = (2u-u^2)(u^2+2u+4)$$

$$= [1-(u-1)^2][(u+1)^2+3] \quad .$$

Also

$$u^2 = -\frac{1}{2}\left[1-(u-1)^2\right] + \frac{1}{2}\left[3+(u+1)^2\right] - 2 \; .$$

Hence,

$$I = -\frac{1}{2}\int\sqrt{\frac{1-(u-1)^2}{3+(u+1)^2}}\,du + \frac{1}{2}\int\sqrt{\frac{3+(u+1)^2}{1-(u-1)^2}}\,du$$

(I.1)

$$-2\int\frac{du}{\sqrt{[1-(u-1)^2][3+(u+1)^2]}}$$

$$= -\frac{1}{2}I_1 + \frac{1}{2}I_2 - 2I_3 \; .$$

§I.1. Algebraic Preliminaries :

Before proceeding with the reduction of the three integrals of (I.1) to standard elliptic form, a few useful results from Whittaker and Watson pages 513, 514, 515 [21] will be reproduced here for completeness of exposition.

§I.1.1. Reduction of the quartic $(8u-u^4)$ to the product of sums of squares :

$$8u-u^4 = (2u-u^2)(u^2+2u+4)$$

$$= S_1 S_2 \quad .$$

Now, let λ be a constant such that $S_1-\lambda S_2$ is a perfect square, that is,

$$\begin{aligned} S_1 - \lambda S_2 &= 2u-u^2 - \lambda(u^2+2u+4) \\ &= (-1-\lambda)u^2+2(1-\lambda)u-4\lambda \end{aligned} \qquad (I.2)$$

is a perfect square. Thus,

$$4(1-\lambda)^2 - 16\lambda(1+\lambda) = 0 \,.$$

Therefore,

$$\lambda_1 = -1+\tfrac{2}{3}\sqrt{3} \quad , \quad \lambda_2 = -1-\tfrac{2}{3}\sqrt{3} \,.$$

By hypothesis there exist constants α and β such that

$$S_1 - \lambda_1 S_2 = -(1+\lambda_1)(u-\beta)^2,$$

$$S_1 - \lambda_2 S_2 = -(1+\lambda_2)(u-\alpha)^2 .$$

Solving these last two equations for S_1 and S_2 we have

$$S_1 = \frac{(1+\lambda_2)\lambda_1}{\lambda_2-\lambda_1}(u-\alpha)^2 - \frac{(1+\lambda_1)\lambda_2}{\lambda_2-\lambda_1}(u-\beta)^2,$$

(I.3)

$$S_2 = \frac{(1+\lambda_2)}{\lambda_2-\lambda_1}(u-\alpha)^2 - \frac{(1+\lambda_1)}{\lambda_2-\lambda_1}(u-\beta)^2 .$$

Writing (I.2) in the form

$$S_1 - \lambda S_2 = -(1+\lambda)\left[u^2 - \frac{2(1-\lambda)}{(1+\lambda)} + \frac{4\lambda}{(1+\lambda)}\right]$$

and substituting the values for λ we have

$$\beta = \sqrt{3} - 1$$

$$\alpha = -\sqrt{3} - 1 .$$

Hence,

$$S_1 = \frac{(2\sqrt{3}-3)\left[u-(-\sqrt{3}-1)\right]^2-(2\sqrt{3}+3)\left[u-(\sqrt{3}-1)\right]^2}{6}$$

(I.4)

$$S_2 = \frac{\left[u-(-\sqrt{3}-1)\right]^2+\left[u-(\sqrt{3}-1)\right]^2}{2} .$$

Let

$$A_1 = \frac{2\sqrt{3}-3}{6}, \quad B_1 = \frac{-(2\sqrt{3}+3)}{6},$$

$$A_2 = \frac{1}{2}, \quad B_2 = \frac{1}{2} .$$

Then (I.4) becomes

$$S_1 = A_1(u-\alpha)^2 + B_1(u-\beta)^2,$$

(I.4a)

$$S_2 = A_2(u-\alpha)^2 + B_2(u-\beta)^2 .$$

§I.1.2 <u>Change of variable</u>:

Let

$$w = \frac{u-\beta}{u-\alpha} .$$

Then,

$$u = \frac{\beta-\alpha w}{1-w}, \quad u-\beta = \frac{(\beta-\alpha)w}{1-w}, \quad u-\alpha = \frac{\beta-\alpha}{1-w},$$

$$du = \frac{\beta-\alpha}{(1-w)^2}\,dw .$$

Hence,

$$S_1 = A_1(u-\alpha)^2 + B_1(u-\beta)^2$$

$$= \left(\frac{\beta-\alpha}{1-w}\right)^2 (A_1 + B_1 w^2) \quad ,$$

$$S_2 = A_2(u-\alpha)^2 + B_2(u-\beta)^2$$

$$= \left(\frac{\beta-\alpha}{1-w}\right)^2 (A_2 + B_2 w^2) \quad .$$

With

$$Q_1 = (A_1 + B_1 w^2) \quad , \quad Q_2 = (A_2 + B_2 w^2)$$

we have

$$S_1 = \left(\frac{\beta-\alpha}{1-w}\right)^2 Q_1 \quad ,$$

(I.4b)

$$S_2 = \left(\frac{\beta-\alpha}{1-w}\right)^2 Q_2 \quad .$$

Using (I.4b) in (I.1) we have

$$I_1 = \int \sqrt{\frac{S_1}{S_2}} \; du$$

$$= \int \frac{(\beta-\alpha)}{(1-w)} \frac{(1-w)}{(\beta-\alpha)} \frac{Q_1^{1/2}}{Q_2^{1/2}} \frac{(\beta-\alpha)}{(1-w)^2} \; dw$$

$$= (\beta-\alpha) \int \frac{(A_1 + B_1 w^2)}{Q_1^{1/2} Q_2^{1/2} (1-w)^2} \; dw$$

$$I_1 = (\beta-\alpha)\int \frac{(A_1+B_1 w^2)(1+w)^2}{Q_1^{1/2}\,Q_2^{1/2}\,(1-w^2)^2}\,dw \,.$$

Thus,

$$I_1 = (\beta-\alpha)\int \frac{1}{Q_1^{1/2}Q_2^{1/2}}\left[\frac{2(A_1+B_1)}{(1-w^2)^2} - \frac{(A_1+3B_1)}{(1-w^2)} + B_1\right]dw$$

$$+\,2(\beta-\alpha)\int \frac{(A_1+B_1 w^2)w}{(1-w^2)^2 Q_1^{1/2}Q_2^{1/2}}\,dw \,.$$

Similarly we can show that

$$I_2 = (\beta-\alpha)\int \frac{1}{Q_1^{1/2}Q_2^{1/2}}\left[\frac{2(A_2+B_2)}{(1-w^2)^2} - \frac{(A_2+3B_2)}{(1-w^2)} + B_2\right]dw$$

$$+\,2(\beta-\alpha)\int \frac{(A_2+B_2 w^2)w}{(1-w^2)^2 Q_1^{1/2}Q_2^{1/2}}\,dw \,.$$

Hence,

$$\text{(I.5)}\quad -\tfrac{1}{2}I_1+\tfrac{1}{2}I_2 = \frac{(\beta-\alpha)}{2}\int \frac{1}{Q_1^{1/2}Q_2^{1/2}}\left[B_2-B_1 - \frac{A_2+3B_2-A_1-3B_1}{(1-w^2)} + \frac{2(A_2+B_2-A_1-B_1)}{(1-w^2)^2}\right]dw$$

$$+\,(\beta-\alpha)\int \frac{w\,(A_2+B_2 w^2-A_1-B_1 w^2)}{(1-w^2)^2\,Q_1^{1/2}\,Q_2^{1/2}}\,dw \,.$$

Finally,

$$\text{(I.5a)}\qquad I_3 = \frac{1}{(\beta-\alpha)}\int \frac{dw}{Q_1^{1/2}Q_2^{1/2}} \,.$$

§I.1.3. A reduction formula:

We wish to reduce integrals of the form

$$\int \frac{dw}{(1+nw^2)^2\, Q_1^{1/2}\, Q_2^{1/2}}\,.$$

Now,

$$\text{(I.6)}\quad \frac{d}{dw}\left[\frac{w}{(1+nw^2)}\, Q_1^{1/2} Q_2^{1/2}\right] = \frac{1}{(1+nw^2)} Q_1^{1/2} Q_2^{1/2} - \frac{2nw^2}{(1+nw^2)^2} Q_1^{1/2} Q_2^{1/2}$$

$$+ \frac{B_1 w^2}{1+nw^2}\, \frac{Q_2^{1/2}}{Q_1^{1/2}} + \frac{B_2 w^2}{1+nw^2}\, \frac{Q_1^{1/2}}{Q_2^{1/2}}$$

$$= \frac{1}{(1+nw^2) Q_1^{1/2} Q_2^{1/2}}\left[Q_1 Q_2 - \frac{2nw^2}{1+nw^2} Q_1 Q_2 + B_1 w^2 Q_2 + B_2 w^2 Q_1\right].$$

Since

$$w^2 = \frac{1}{n}(1+nw^2-1),$$

then

$$w^4 = \frac{1}{n^2}\left[(1+nw^2)^2 - 2(1+nw^2) + 1\right]$$

and

$$w^6 = \frac{1}{n^3}\left[(1+nw^2)^3 - 3(1+nw^2)^2 + 3(1+nw^2) - 1\right].$$

Using these last results we obtain

$$\frac{Q_1 Q_2}{1+nw^2} = (1+nw^2)\frac{B_1 B_2}{n^2} + \left[-\frac{2B_1B_2}{n^2} + \frac{A_1B_2+A_2B_1}{n}\right]$$

$$+\frac{1}{1+nw^2}\left[\frac{B_1B_2}{n^2} - \frac{A_1B_2+A_2B_1}{n} + A_1A_2\right],$$

$$-\frac{2nw^2\, Q_1Q_2}{(1+nw^2)^2} = -\frac{2Q_1Q_2}{1+nw^2} + \frac{2Q_1Q_2}{(1+nw^2)^2} \; .$$

Thus,

$$\frac{Q_1Q_2}{1+nw^2} - \frac{2nw^2Q_1Q_2}{(1+nw^2)^2} = -(1+nw^2)\frac{B_1B_2}{n^2} + \left[\frac{4B_1B_2}{n^2} - \frac{A_1B_2+B_2A_1}{n}\right]$$

$$+\frac{1}{1+nw^2}\left[-\frac{5B_1B_2}{n^2} + \frac{3(A_1B_2+A_2B_1)}{n} - A_1A_2\right]$$

$$+\frac{1}{(1+nw^2)^2}\left[\frac{2B_1B_2}{n^2} - \frac{2(A_1B_2+A_2B_1)}{n} + 2A_1A_2\right].$$

Also,

$$\frac{B_1w^2Q_2}{1+nw^2} = \frac{B_1B_2w^4 + B_1A_2w^2}{1+nw^2} \; .$$

Hence,

$$\frac{B_1w^2Q_2}{1+nw^2} + \frac{B_2w^2Q_1}{1+nw^2} = 2(1+nw^2)\frac{B_1B_2}{n^2} + \left[-\frac{4B_1B_2}{n^2} + \frac{B_1A_2+A_1B_2}{n}\right] + \frac{1}{1+nw^2}\left[\frac{2B_1B_2}{n^2} - \frac{A_1B_2+B_1A_2}{n}\right].$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

Thus, (I.6) becomes

$$\frac{d}{dw}\left[\frac{w\,Q_1^{1/2}Q_2^{1/2}}{(1+nw^2)}\right] = \frac{2}{(1+nw^2)^2 Q_1^{1/2}Q_2^{1/2}}\left[\frac{B_1B_2}{n^2} - \frac{A_1B_2+A_2B_1}{n} + A_1A_2\right]$$

$$+\frac{1}{(1+nw^2)Q_1^{1/2}Q_2^{1/2}}\left[-\frac{3B_1B_2}{n^2} + \frac{2(A_1B_2+A_2B_1)}{n} - A_1A_2\right] + \frac{(1+nw^2)B_1B_2}{Q_1^{1/2}Q_2^{1/2}\,n^2}\,.$$

Hence, upon integrating with respect to w we have

$$\text{(I.7)}\qquad 2\left[\frac{B_1B_2}{n^2} - \frac{A_1B_2+A_2B_1}{n} + A_1A_2\right]\int\frac{dw}{(1+nw^2)^2Q_1^{1/2}Q_2^{1/2}}$$

$$= \left[\frac{3B_1B_2}{n^2} - \frac{2(A_1B_2+A_2B_1)}{n} + A_1A_2\right]\int\frac{dw}{(1+nw^2)Q_1^{1/2}Q_2^{1/2}}$$

$$-\frac{B_1B_2}{n^2}\int\frac{(1+nw^2)\,dw}{Q_1^{1/2}Q_2^{1/2}} + \left[\frac{w\,Q_1^{1/2}Q_2^{1/2}}{1+nw^2}\right]_{w_1}^{w},$$

where

$$w_1 = \frac{u_1-\beta}{u_1-\alpha}\,,\quad u_1 \text{ as in equation (4.9)}.$$

§I.2. Completion of the Reduction of $\int\frac{u^2\,du}{\sqrt{8u-u^4}}$:

§I.2.1. Specialization:

With

$$A_1 = \frac{2\sqrt{3}-3}{6}, \qquad B_1 = \frac{-(2\sqrt{3}+3)}{6},$$

$$A_2 = \frac{1}{2}, \qquad B_2 = \frac{1}{2}, \quad n = -1$$

$$\beta = \sqrt{3} - 1, \qquad \alpha = -\sqrt{3} - 1$$

let us evaluate the coefficients of the terms in (I.5) and (I.7). Thus, (I.5) becomes

$$\text{(I.8)} \quad -\tfrac{1}{2} I_1 + \tfrac{1}{2} I_2 = \sqrt{3} \int \frac{1}{Q_1^{1/2} Q_2^{1/2}} \left[\left(\frac{6+2\sqrt{3}}{6} \right) - \frac{2}{3} \left(\frac{6+\sqrt{3}}{1-w^2} \right) + \frac{4}{(1-w^2)^2} \right] dw$$

$$+ 2\sqrt{3} \int \frac{[(6-2\sqrt{3}) + w^2(6+2\sqrt{3})]w}{6(1-w^2)^2 Q_1^{1/2} Q_2^{1/2}} dw$$

and (I.7) becomes

$$\text{(I.9)} \quad \int \frac{dw}{(1-w^2)^2 Q_1^{1/2} Q_2^{1/2}} = \frac{6+\sqrt{3}}{6} \int \frac{dw}{(1-w^2) Q_1^{1/2} Q_2^{1/2}} - \frac{3+2\sqrt{3}}{24} \int \frac{(1-w^2)}{Q_1^{1/2} Q_2^{1/2}} dw$$

$$- \frac{1}{2} \left[\frac{w\, Q_1^{1/2} Q_2^{1/2}}{1-w^2} \right]_{w_1}^{w} .$$

§I.2.2. Application of (I.9) to (I.8):

Let us write (I.8) in the form

$$-\tfrac{1}{2} I_1 + \tfrac{1}{2} I_2 = 4\sqrt{3} \int \frac{dw}{(1-w^2)^2 Q_1^{1/2} Q_2^{1/2}} + (\sqrt{3}+1) \int \frac{dw}{Q_1^{1/2} Q_2^{1/2}} - 2(2\sqrt{3}+1) \int \frac{dw}{(1-w^2) Q_1^{1/2} Q_2^{1/2}}$$

$$+ 2 \int \frac{[(\sqrt{3}-1) + w^2(\sqrt{3}+1)]w}{(1-w^2)^2 Q_1^{1/2} Q_2^{1/2}} dw .$$

Apply (I.9) to the first integral on the right to get

$$-\tfrac{1}{2}I_1 + \tfrac{1}{2}I_2 = (\sqrt{3}+1)\int \frac{dw}{Q_1^{1/2}Q_2^{1/2}} - \frac{\sqrt{3}+2}{2}\int \frac{(1-w^2)\,dw}{Q_1^{1/2}Q_2^{1/2}}$$

$$+ 2\int \frac{[(\sqrt{3}-1)+w^2(\sqrt{3}+1)]w}{(1-w^2)^2 Q_1^{1/2}Q_2^{1/2}}\,dw \;-\; 2\sqrt{3}\left[\frac{w\,Q_1^{1/2}Q_2^{1/2}}{1-w^2}\right]_{w_1}^{w} .$$

Now, from (I.5a) we get

$$-2I_3 = -\frac{\sqrt{3}}{3}\int \frac{dw}{Q_1^{1/2}Q_2^{1/2}} .$$

Hence,

$$I = -\tfrac{1}{2}I_1 + \tfrac{1}{2}I_2 - 2I_3$$

$$= \frac{2\sqrt{3}+3}{3}\int \frac{dw}{Q_1^{1/2}Q_2^{1/2}} - \frac{\sqrt{3}+2}{2}\int \frac{(1-w^2)\,dw}{Q_1^{1/2}Q_2^{1/2}} + 2\int \frac{[(\sqrt{3}-1)+w^2(\sqrt{3}+1)]w}{(1-w^2)^2 Q_1^{1/2}Q_2^{1/2}}\,dw$$

$$-2\sqrt{3}\left[\frac{w\,Q_1^{1/2}Q_2^{1/2}}{1-w^2}\right]_{w_1}^{w} .$$

$$\text{(I.10)}\quad I = \left(\frac{2\sqrt{3}+3}{3}\right)I_4 - \left(\frac{\sqrt{3}+2}{2}\right)I_5 + 2I_6 - f_1(w) .$$

Note that I_6 can be evaluated in terms of elementary functions.

§ I.2.3. The integral I_4 :

$$I_4 = \int \frac{dw}{Q_1^{1/2} Q_2^{1/2}} \quad ,$$

where

$$Q_1 = \frac{2\sqrt{3}-3}{6} - \left(\frac{2\sqrt{3}+3}{6}\right) w^2$$

$$= \frac{2\sqrt{3}-3}{6}\left[1 - \frac{(2\sqrt{3}+3)^2}{3} w^2\right] ,$$

$$Q_2 = \frac{1}{2}(1+w^2) \quad .$$

So,

$$I_4 = 2(2\sqrt{3}+3)^{1/2} \int \frac{dw}{(1+w^2)^{1/2}(1-\gamma^2 w^2)^{1/2}} \quad ,$$

where

$$\gamma^2 = \frac{(2\sqrt{3}+3)^2}{3} \quad .$$

Let

$$w = \frac{v}{\sqrt{1-v^2}} \quad .$$

Then I_4 becomes

$$I_4 = 2(2\sqrt{3}+3)^{1/2} \int \frac{dv}{\sqrt{(1-v^2)[1-(1+\gamma^2)v^2]}} \quad .$$

Finally, with

$$1+\gamma^2 = \frac{1}{k^2} \quad , \quad v = k\sigma$$

we have

$$\text{(I.11)} \qquad I_4 = 2(2\sqrt{3}+3)^{1/2}\, k \int \frac{d\sigma}{\sqrt{(1-k^2\sigma^2)(1-\sigma^2)}}$$

which is the elliptic integral of the first kind, in standard form.

§ I.2.4. The integral I_5 :

Now,

$$I_5 = \int \frac{(1-w^2)\,dw}{Q_1^{1/2}\,Q_2^{1/2}}$$

$$= 2(2\sqrt{3}+3)^{1/2} \int \frac{(1-w^2)\,dw}{\sqrt{(1+w^2)(1-\gamma^2 w^2)}} \, .$$

Let

$$w = \frac{v}{\sqrt{1-v^2}} \, .$$

Then

$$1-w^2 = 1 - \frac{v^2}{1-v^2}$$

$$= \frac{1-2v^2}{1-v^2}$$

$$= 2 - \frac{1}{1-v^2} \, .$$

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

Thus,

$$I_5 = 2(2\sqrt{3}+3)^{1/2}\int \frac{2\,dv}{\sqrt{(1-v^2)[1-(1+\gamma^2)v^2]}}$$

$$-2(2\sqrt{3}+3)^{1/2}\int \frac{dv}{(1-v^2)\sqrt{(1-v^2)[1-(1+\gamma^2)v^2]}} .$$

With $1+\gamma^2 = 1/k^2$, $v = k\sigma$ we obtain

$$\text{(I.12)} \quad I_5 = 4(2\sqrt{3}+3)^{1/2} k \int \frac{d\sigma}{\sqrt{(1-k^2\sigma^2)(1-\sigma^2)}}$$

$$-2(2\sqrt{3}+3)^{1/2} k \int \frac{d\sigma}{(1-k^2\sigma^2)\sqrt{(1-k^2\sigma^2)(1-\sigma^2)}} .$$

The last integral above is an elliptic integral of the second kind, in standard form.

§ I.2.5. The integral I_6 :

Now,

$$I_6 = \int \frac{[(\sqrt{3}-1)+(\sqrt{3}+1)w^2]\,w}{(1-w^2)^2\,Q_1^{1/2}\,Q_2^{1/2}}\,dw .$$

It can be show that

$$\text{(I.13)} \qquad I_6 = \frac{\sqrt{3}\;Q_1^{1/2}\,Q_2^{1/2}}{1-w^2} .$$

[illegible]

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

[illegible]

§I.2.6. The final result:

With (I.11), (I.12), (I.13) in (I.10) we have

$$I = -\frac{2}{3}k\,(72+42\sqrt{3})^{1/2}\int\frac{d\sigma}{\sqrt{(1-k^2\sigma^2)(1-\sigma^2)}}$$

$$+(26\sqrt{3}+45)^{1/2}k\int\frac{d\sigma}{(1-k^2\sigma^2)\sqrt{(1-k^2\sigma^2)(1-\sigma^2)}} - \left[\frac{2\sqrt{3}\,Q_1^{1/2}\,Q_2^{1/2}}{1-w}\right]_{w_1}^{w}.$$

Hence, we have

$$\int_{u_1}^{X^2}\frac{u^2\,du}{\sqrt{8u-u^4}} = (26\sqrt{3}+45)^{1/2}k\int_{\sigma_1}^{\sigma}\frac{d\sigma}{(1-k^2\sigma^2)\sqrt{(1-k^2\sigma^2)(1-\sigma^2)}}$$

$$-\frac{2}{3}k\,(72+42\sqrt{3})^{1/2}\int_{\sigma_1}^{\sigma}\frac{d\sigma}{\sqrt{(1-k^2\sigma^2)(1-\sigma^2)}} - 2\sqrt{3}\left[\frac{Q_1^{1/2}\,Q_2^{1/2}}{1-w}\right]_{w_1}^{w},$$

where

$$w = \frac{X^2-\eta(\sqrt{3}-1)}{X^2+\eta(\sqrt{3}+1)},$$

$$\sigma = \frac{X^2-\eta(\sqrt{3}-1)}{k\sqrt{2}\left[X^4+2\eta X^2+4\eta^2\right]^{1/2}},$$

$$k = \frac{\sqrt{2}}{4}(\sqrt{3}-1).$$

APPENDIX II

THE EIGENQUANTITIES OF D_k^2

Note that D_k^2 is a positive definite, symmetric matrix and so has k real, distinct eigenvalues and k linearly independent eigencolumns. The eigenrows are the transpose of the eigencolumns.

We wish to find scalars (λ) and column vectors (s) such that

$$D_k^2 s = \lambda s. \tag{II.1}$$

Denote the elements of s by

$$s_p, \quad p = 1, 2, \ldots, k.$$

Equating elements in the pth row of the above equation we obtain

$$-s_{p-1} + 2s_p - s_{p+1} = \lambda s_p. \tag{II.2}$$

This relation holds for all $p = 1, 2, \ldots, k$ if we define

$$s_0 = s_{k+1} = 0.$$

We can rewrite (II.2) as

[illegible]

[illegible]

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

(II.3) $$s_{p+1} - (2-\lambda) s_p + s_{p-1} = 0 .$$

Now, we know that

$$\lambda \geq 0$$

because D_k^2 is positive definite. It is also known that

$$|\lambda| \leq 4 .$$

Hence,

$$0 \leq \lambda \leq 4$$

and we can let

$$\lambda = 2(1 - \cos \theta).$$

Substituting this in (II.3) we have

(II.4) $$s_{p+1} - (2 \cos \theta) s_p + s_{p-1} = 0.$$

This is a difference equation for s. Let us try the solution

$$s_p = t^p,$$

where t is a constant. Substituting this in

(II.4) and dividing through by t^{p-1} we get

$$t^2 - 2t\cos\theta + 1 = 0$$

which has the solutions

$$t = e^{i\theta}, \quad t = e^{-i\theta}, \quad i = \sqrt{-1}\ .$$

The general solution to (II.4) is

(II.5) $$s_p = Ae^{ip\theta} + Be^{-ip\theta},$$

where A and B are constants. Since $s_0 = 0$

$$A = -B.$$

Hence, (II.5) becomes

(II.5a) $$s_p = A\sin(p\theta).$$

Since $s_{k+1} = 0$ we have

$$\sin[(k+1)\theta] = 0$$

and so

$$\theta = \frac{l\pi}{k+1}, \quad l = 1, 2, \ldots\ .$$

Therefore,

$$\lambda_\ell = 2\left[1 - \cos\left(\frac{\ell\pi}{k+1}\right)\right], \quad \ell = 1, 2, \ldots, k.$$

If we now denote the elements of the column vector s_ℓ by $s_{p\ell}$ we have

$$s_{p\ell} = A \sin\left(\frac{p\ell\pi}{k+1}\right), \quad p = 1, 2, \ldots, k.$$

The constant A is determined by the orthogonality relation

$$s_\ell^t s_q = \delta_{\ell q} \quad ,$$

that is,

$$A^2 \sum_{p=1}^{k} \sin^2\left(\frac{p\ell\pi}{k+1}\right) = 1 .$$

Thus

$$\frac{1}{A^2} = \sum_{p=1}^{k} \sin^2\left(\frac{p\ell\pi}{k+1}\right)$$

$$= \frac{1}{2} \sum_{p=1}^{k} \left[1 - \cos\left(\frac{2p\ell\pi}{k+1}\right)\right] .$$

To simplify notation let

$$\frac{2\ell\pi}{k+1} = \theta .$$

Thus

$$\frac{1}{A^2} = \frac{1}{2}\left[k - \sum_{p=1}^{k} \cos p\theta \right]$$

$$= \frac{1}{2}\left[k - Re \sum_{p=1}^{k} e^{ip\theta} \right]$$

$$= \frac{1}{2}\left[k - Re \left\{ \frac{e^{i\theta}(1 - e^{ik\theta})}{1 - e^{i\theta}} \right\} \right]$$

$$= \frac{1}{2}\left[k - Re \left\{ \frac{e^{i\theta} - e^{i(k+1)\theta}}{1 - e^{i\theta}} \right\} \right].$$

Now,

$$e^{i(k+1)\theta} = e^{2\pi i \ell} = 1 .$$

Hence,

$$\frac{1}{A^2} = \frac{k+1}{2}$$

and so

$$s_{p\ell} = \sqrt{\frac{2}{k+1}} \quad \sin\left(\frac{p\ell\pi}{k+1}\right).$$

In summary, the matrix D_k^2 has k eigenvalues

$$\lambda_\ell = 2\left[1 - \cos\left(\frac{\ell\pi}{k+1}\right)\right]$$

and k linearly independent eigencolumns s_ℓ with elements

$$s_{p\ell} = \sqrt{\frac{2}{k+1}} \quad \sin\left(\frac{p\ell\pi}{k+1}\right) ,$$

where

$$\ell, p = 1, 2, \ldots, k .$$

The eigenrows of D_k^2 are transposes of the eigencolumns.